Computer Methods in Applied Mechanics and Engineering 429 (2024) 117192

Contents lists available at ScienceDirect

Comput. Methods Appl. Mech. Engrg.

journal homepage: www.elsevier.com/locate/cma

FE-LSTM: A hybrid approach to accelerate multiscale simulations of architectured materials using Recurrent Neural Networks and Finite Element Analysis

Aymen Danoun [a], Etienne Prulière [b], Yves Chemisky [a,*]

[a] Univ. Bordeaux, CNRS, Bordeaux INP, I2M, UMR 5295, Esplanade des Arts et Métiers, Talence, 33400, France
[b] Arts et Metiers Institute of Technology, CNRS, Bordeaux INP, Hesam Universite, I2M, UMR 5295, Esplanade des Arts et Métiers, Talence, 33400, France

ARTICLE INFO

Keywords:
Recurrent Neural Network (RNN)
Architectured materials
Multi-scale modeling
Homogenization
Artificial Intelligence
Finite Element Analysis

ABSTRACT

In the present work, a novel modeling strategy to accelerate multi-scale simulations of heterogeneous materials using deep neural networks is developed. This approach, called FE-LSTM, consists of combining the finite element method and LSTM recurrent neural networks to solve multiscale problems. In contrast to FE^2 method, the finite element resolution of the microscopic problems is no longer required within FE-LSTM framework, the computation of RVE homogenized response is directly predicted by an LSTM trained on a database of offline micro simulations. The usefulness of such approach lies in the fact that once the RNN training process is performed, FE-LSTM can be applied to simulate nonlinear behaviors of any heterogeneous structure having the same microstructure used in training, considering a different Boundary Value Problem to solve. To assess the validity and reliability of the developed approach, FE-LSTM has been evaluated on several 3D architectured structures from rather simple to complex geometries under proportional and non proportional loading conditions. In terms of execution time, it has been found that FE-LSTM shows a speedup factor of nearly 40 000 as compared with legacy two-scale implementation of the FE^2 method, while up to date optimized FE^2 methods shows a speed up factor around 1000. Furthermore, RAM memory saving factor allow complex structures simulations to be conducted on desktop computers without requiring HPC clusters with specific additions to RAM memory.

1. Introduction

The use of heterogeneous materials has seen an increase of interest in the last decades for the design of advanced engineering applications, including renewable energy, healthcare and aerospace. For instance, the emergence of cellular architectured materials is opening new possibilities thanks to their high strength-to-weight ratio and significant energy absorption capabilities Specifically, lattice-type and triply-periodic minimal surfaces based architectured structures can be employed for many applications, and their tunable mechanical properties is particularly suitable in biomechanics to conceive orthopedic implants [1]. Through the optimization of microstructures topology, architectured prostheses can be designed to improve the mechanical compatibility with human bones. Moreover, they can be employed to ensure other import functionalities, such as nutriment transport and bone cell colonization.

However, modeling the thermo-mechanical response of heterogeneous structures can be very challenging due to many factors: First, the mechanical behavior resulting from the microstructures is very complex to identify as multiple physical mechanisms are

* Corresponding author.
E-mail address: yves.chemisky@u-bordeaux.fr (Y. Chemisky).

https://doi.org/10.1016/j.cma.2024.117192
Received 11 February 2024; Received in revised form 27 May 2024; Accepted 21 June 2024
Available online 5 July 2024

usually involved at different length scales. Local material behaviors are particularly governed by strongly non-linear dissipative phenomena that requires a proper identification of the constitutive laws. In addition, multiscale modeling strategies are necessarily required to have a better understanding of the effect of the microstructure on the macroscopic response. Consequently, the evaluation of heterogeneous structures mechanical behaviors requires the development of advanced and powerful numerical tools, capable of taking into consideration the micro-macro interaction in real time and predicting accurately the resulting global response. Among these computational approaches, the multi level finite element method FE^2 [2,3] have been extensively developed in the literature for multiscale structural analysis. This full field approach based on periodic homogenization theory, consist on attributing a Representative Volume Element (RVE) to each Gauss integration point. The estimation of the overall response is then obtained by solving micro and macro problems simultaneously through localization and homogenization principles. Although FE^2 remains more advantageous in terms of execution time compared to direct FE simulations on fully meshed heterogeneous structures, it is clear that such modeling strategy also suffers from numerous limitations given the following reasons: as described in the above mentioned FE^2 procedure, nonlinear problems are solved simultaneously on both micro and macro scales using Newton–Raphson iterative schemes. This process is commonly time consuming due to the treatment of material non-linearities in each microscopic computation and at each iteration [4,5]. In addition, other considerations may increase the computational time including the number of Gauss integration points, the complexity of loading conditions, the number of increments and the meshing resolution of the RVE and the macrostructure. Therefore, all these considerations may sometimes hinder FE^2 approach from being efficiently applicable to structural analysis.

Note that several efforts have been made to optimize such schemes, such as the monolithic resolution [6] which has shown a speed-up factor of about 2.5. Also, an unsupervised machine learning-based clustering approach has been developed by [7], which shows-up speed up factors of about 20. Also, the monolithic solution algorithm has been combined with reduced order modeling (ROM) and the empirical cubature method (ECM) in [8] which allows speed-up factors of about 1000. According to these recent improvements, the reader should be warned that the speed-up factors that are presented in Section 4 are always presented with reference to a legacy approach, since those calculations are performed on the same finite element solver (fedoo).

FE^2 is not the only multi-scale model that is able to consider the computation of mechanical fields without specific assumptions . For instance, numerous numerical methods based on the Fast Fourier Transform (FFT) have been developed over the years following the seminal work of *Moulinec and Suquet* [9]. Reduced-order models (ROM) based on this methods have been already successfully applied by *Michel an Suquet* [10], namely the nonuniform transformation field analysis (NTFA), with several enhancements over the years , summarized in [11]. In the present work, we seek to study an alternative method based on the use of Artificial Intelligence-based models to evaluate their capability in terms of computational time breakdown

Artificial Intelligence (AI) is one of the most prominent technologies to provide a broad and fast access to legacy computationally intensive simulations through Machine Learning (ML) and Deep Learning (DL). In material science, the recent developments in experimental measurement techniques have enabled the generation of large databases of material behaviors that can be used for constitutive models identification. On the other hand, the significant advances in computational capabilities have facilitated the production of extensive quantities of numerical simulations data. From this perspective, the employment of machine learning and deep learning has considerably attracted the interest of the material science community. For instance, the use of ANN as surrogates has been proven to be a reliable tool for several applications including structural analysis, constitutive modeling and multi-scale modeling. The first application of neural networks in material science was in 1991 where *Ghaboussi et al.* [12] used a Multi Layer Perceptron (MLP) or a Feed Forward Neural Network (FFNN) to model the mechanical behavior of concrete under monotonic biaxial loading and compressive uniaxial loading. *Furukawa et al.* [13] proposed an implicit viscoplastic constitutive model using a MLP that takes as inputs the current viscoplastic strain, internal variables and current stress, then predict the current rates of change of the viscoplastic strain and material internal variables. Another extension of ANN to rate dependant materials modeling has been proposed by [14], this study consists of the implementation of a neural network, trained to capture viscoelasticity, in a finite element code and then applied to structural analysis of concrete. Using Neural Networks back-propagation algorithm as a replacement for numerical procedures, *Waszczyszyn et al.* [15] treated several problems regarding elastoplasticity such as the bending analysis of elastoplastic beams and detection of damage in steel beams.*Zhang and Mohr* [16] reformulated the standard return-mapping algorithm by substituting the non linear procedures (used to compute stress–strain relationships and elastoplastic tangent matrix) with a feed forward neural network. This method has been employed for the prediction of Von Mises plasticity with isotropic hardening. In a recent study, while *Ali et al.* [17] applied a MLP architecture to predict stress–strain responses and texture evolution of polycrystalline metals under shear and tension loading, the computation time saving was estimated at 10 000 by adopting this approach. Similarly, *Shen et al.* [18] used an ANN model to predict plastic yield surfaces for porous materials with different loading states and different porosity coefficients. *Settgast et al.* [19] also applied a FFNN as a constitutive model to describe deformation behavior of open-cell foams under proportional loading paths, the computational time saving factor was also estimated at 10 000 during this study. According to such recent studies, artificial neural networks have proven to be reliable and effective methods for engineering material design due to their capabilities to accurately predict target material properties. For instance, a feed forward neural network has been employed by [20] to capture the associated non-linear composition–property relationships of oxide glasses. In another study conducted by *Liu et al.* [21], ANN have been used in a fracture mechanics framework to predict the toughness of polysilicon materials. The high predictive capability of deep neural networks, even when trained on small datasets, has been highlighted by *Ouyang et al.* [22] to predict concrete's strength as a function of the material mixture proportions. They have shown that ANN are able to quickly and successfully learn input–output relationships compared to other machine learning algorithms.

Recurrent Neural Network (RNN) and in particular Long Short-Term Memory (LSTM) networks have received a particular interest in the recent years to describe the response of materials and structures exhibiting inelastic strains. [23] has shown that this class of network is well suited to represent viscoplasticity considering strain rate and temperature history dependence. [24] has shown that a thermodynamically consistent LSTM network can represent with a great accuracy the response of elastoplastic materials considering both isotropic and kinematical hardening. [25] have developed an LSTM surrogate model to predict the heterogeneous deformation behavior of dual phase microstructures. It is shown that such architecture is able to predict the mechanical response and the strain partitioning in the dual phase microstructures, considering training based on Finite Element simulation of dual phase microstructures equipped with a dislocation based J_2 plasticity. [26] have also proposed an LSTM surrogate model adapted to large elastoplastic FE simulations that combines LSTM neural network with (proper orthogonal decomposition) POD data reduction. A database is first constructed using commercial FE analysis software. The high-dimensional input data are reduced to low-dimensional POD (proper orthogonal decomposition) coefficient data before being used for training. [27] have investigated the capability of LSTM network to represent the elastic–plastic response of materials. They have found that such networks are able to retrieve the exact elastoplastic response (considering uniaxial response) if the model is piecewise linear, considering validation over complex temporal input signals. In such recent work, the robustness of LSTM has been highlighted as a powerful tool capable of handling path dependent behaviors and accurately capturing complex and highly non-linear behaviors, hence the employment of such architecture to approximate the effective material responses in this multiscale framework.

A few recent contributions have also demonstrated that deep learning methods can be integrated into a computational homogenization framework to accelerate multi-scale simulations. For example, *Le et al.* [28] used a neural network to compute the mechanical behavior of elastic non-linear heterogeneous structures under proportional loading paths. The approach involved replacing the strain energy density of the homogenized material with a response surface predicted by an ANN. The derivation of this effective potential allowed to obtain the macroscopic constitutive law. This method was applied to compute the mechanical response of a multi-scale composite structure. Another example of substituting the energy density of heterogeneous materials by an ANN-based approach was studied by *MinhNguyen et al.* [29]. This method was applied to the computational homogenization of non-linear elastic structures with finite strain. This study demonstrated that the multilayer perceptron, trained on a database of FFT simulations, showed excellent predictive capabilities of the macroscopic potentials. The robustness and reliability of deep learning methods in computational homogenization was also highlighted in another study conducted by *Lu et al.* [30] to characterize the non-linear electric conduction in random graphene-polymer nano-composites. This data driven approach involved using an ANN-based surrogate to construct the effective electrical constitutive law after a training phase on RVE non-linear electrical conduction simulations. By adopting this approach, the online speed up factor was estimated to be 104, thus significantly reducing the computational time compared to regular FE^2 method. However, it is worth mentioning that the conducted studies by several authors [28–30] were limited to non-dissipative materials under proportional loading paths, thus leading to a significant simplification of the problem.

The aim of the present work is to go beyond those assumptions by considering non-linear dissipative materials under multi-axial and non proportional loading paths. For this purpose, Recurrent Neural Networks (RNN) presented in [31] are used instead of classical MLP to capture history dependent behaviors. Contrary to MLPs, RNN are designed to handle sequential data in a recursive manner. This type of architecture have been extensively applied in several tasks such as speech recognition, text generation or machine translation. For material science applications, RNN based architectures has been recently employed for constitutive modeling of plasticity and thermo-visco-plasticity, respectively [32,33]. Within this paper, we propose to develop a hybrid approach FE-LSTM that combines a specific RNN architecture called Long Short Term Memory (LSTM) neural network and finite element analysis to accelerate multi-scale simulation of non-linear dissipative architectured materials under multi-axial and non proportional loading paths. For this purpose, the present work is organized as follows: In Section 2, the main objective is to establish the theoretical foundations behind multi-scale modeling approaches whose principles constitute the core of computational homogenization even when combined with deep neural networks. Section 3 presents first an overview of LSTM neural networks architecture and secondly the design of FE-LSTM: an approach that combines the FE method and LSTM neural networks for architectured structures multi-scale modeling. In Section 4, the entire workflow of FE-LSTM model from database generation, hyperparameters selection, training to validation is detailed. Afterwards, the application of FE-LSTM model to the multi-scale analysis of architectured structures in real life condition is illustrated. The obtained results with FE-LSTM model are compared with the FE^2 approach in terms of accuracy and computational costs savings (execution time, required memory, CPU usage). Finally, some concluding remarks are given in Section 5.

2. Theoretical framework of computational homogenization

2.1. Periodic homogenization for heterogeneous media

Homogenization techniques consists in identifying a global mechanical behavior at the macroscopic scale *i.e* the effective mechanical response given the microstructure characteristics (mechanical properties, geometry, size, orientation, volume fraction, constitutive laws). As shown in Fig. 1, the objective is to define an equivalent homogeneous medium having an identical average mechanical response (stress and strain fields) to the heterogeneous structure.

An heterogeneous medium is considered triply periodic when it is defined by a repeating unit cell representative of the microstructure to fill $\mathcal{R}^3$ space. The theory of periodic homogenization is valid provided that scale separation conditions between the microscale and macro-scale are verified. At the micro-scale, an RVE can be defined as a material volume sufficiently large enough

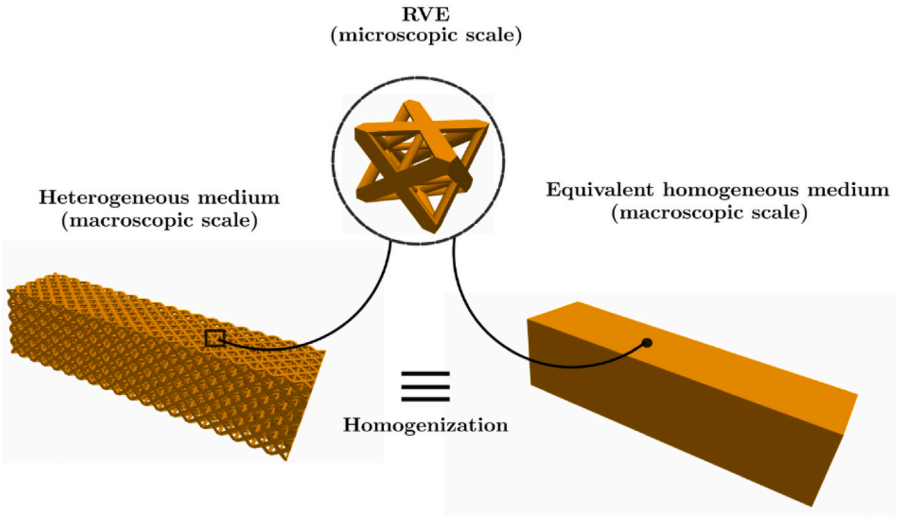

Fig. 1. Schematic representation of the homogenization principles.

to be statistically representative to the microstructure. At the macro-scale, the structure is considered as a virtual homogeneous medium where each macroscopic material point can be associated to a RVE. The assumption of scale separation is valid as long as the characteristic size of the unit cell (microstructure) is several orders of magnitude smaller than the dimensions of the macrostructure.

Assuming that the boundary $\partial\Omega$ of a periodic RVE is subjected to periodicity conditions, meaning that the displacement field $\boldsymbol{u}(\boldsymbol{x})$ of any material point with a position vector $\boldsymbol{x}$ can be expressed by an affine part $\bar{\boldsymbol{\varepsilon}} \cdot \boldsymbol{x}$ and a periodic fluctuating displacement field $\tilde{\boldsymbol{u}}$:

$$\boldsymbol{u}(\boldsymbol{x}) = \bar{\boldsymbol{\varepsilon}} \cdot \boldsymbol{x} + \tilde{\boldsymbol{u}}. \tag{1}$$

The periodic fluctuating quantity $\tilde{\boldsymbol{u}}$ remains the same for each pair of opposite nodes on the RVE boundary. As shown in Fig. 2, the displacement field of two opposite nodes i and j with a position vector $\boldsymbol{x}^+$ and $\boldsymbol{x}^-$ respectively is expressed as:

$$\boldsymbol{u}\left(\boldsymbol{x}^+\right) = \bar{\boldsymbol{\varepsilon}} \cdot \boldsymbol{x}^+ + \tilde{\boldsymbol{u}} \tag{2}$$

$$\boldsymbol{u}\left(\boldsymbol{x}^-\right) = \bar{\boldsymbol{\varepsilon}} \cdot \boldsymbol{x}^- + \tilde{\boldsymbol{u}} \tag{3}$$

In that case, each pair of opposite nodes are linked by the following kinematic equation [34]:

$$\boldsymbol{u}\left(\boldsymbol{x}^+\right) - \boldsymbol{u}\left(\boldsymbol{x}^-\right) = \nabla \boldsymbol{u} \cdot \boldsymbol{\Delta} x, \tag{4}$$

which reduces to Eq. (5) under small strains:

$$\boldsymbol{u}\left(\boldsymbol{x}^+\right) - \boldsymbol{u}\left(\boldsymbol{x}^-\right) = \bar{\boldsymbol{\varepsilon}} \cdot \left(\boldsymbol{x}^+ - \boldsymbol{x}^-\right), \tag{5}$$

Given the aforementioned periodicity conditions, the volume average of the strain inside the RVE is expressed by:

$$\frac{1}{V} \int_{\Omega} \varepsilon \mathrm{d}V = \bar{\varepsilon} \tag{6}$$

Under small strain theory assumptions, average stress and average strain theorems state that the stress and strain averages within the RVE are equal to the applied uniform traction and linear displacement on its boundary respectively, see for instance [35]. Although asymptotic expansion is the most general framework to define the periodic homogenization problem, in the zeroth order periodic homogenization equivalent results are provided for the mechanical problem following the methodology defined by [34] and described above. Considering non-linear response of materials, the macro-scale and unit cell problems are solved simultaneously, using an iterative scheme, referred to as a linearized, incremental formulation. The unit cell problem provides information to compute the macroscopic mechanical tangent modulus, which is required in the macro-scale analysis. The computational framework includes two types of increments of a quantity x: (i) increment in time, denoted by Δx and (ii) increment during the Newton–Raphson scheme in the linearized equations, denoted by the symbol δx. The equations of the unit cell problem can therefore be written in the linearize, incremental form as [35]:

$$\delta\boldsymbol{\epsilon} = \delta\bar{\boldsymbol{\varepsilon}} + \mathrm{sym}\nabla\tilde{\mathbf{u}}\,. \tag{7}$$

The localization equation can be written as:

$$\delta\boldsymbol{\epsilon} = \boldsymbol{\mathcal{A}} : \delta\bar{\boldsymbol{\varepsilon}}\,, \tag{8}$$

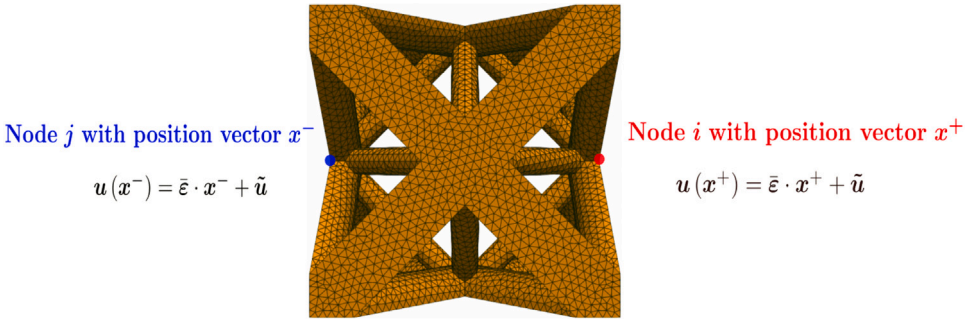

Fig. 2. Schematic representation of two opposite nodes at the boundary of an octet-truss RVE.

Table 1
Equations that define the mechanical problem at the micro and macro-scale.

Control equations	Macro-scale problem	Micro-scale problem
Equilibrium	$\mathrm{div}\,\bar{\boldsymbol{\sigma}} + \bar{\boldsymbol{b}} = 0$	$\mathrm{div}\,\boldsymbol{\sigma} = 0$
Kinematics	$\bar{\varepsilon} = \frac{1}{2}\left(\nabla\bar{u} + {}^t\boldsymbol{\nabla}\bar{u}\right)$	$\bar{\varepsilon} = \frac{1}{2}\left(\nabla u + {}^t\boldsymbol{\nabla} u\right)$
Stress-Strain incremental constitutive incremental constitutive law	$\Delta\bar{\sigma} = \bar{\boldsymbol{C}}^{\tan}\Delta\bar{\varepsilon}$	$\Delta\sigma = \bar{\boldsymbol{C}}^{\tan} : \Delta\varepsilon$

Table 2
Scale coupling set of equations.

Control equations	Scale coupling
Periodic boundary conditions	$\Delta u^+ - \Delta u^- = \Delta\bar{\varepsilon} \cdot \left(x^+ - x^-\right)$
Homogenization	$\bar{\sigma} = \langle\boldsymbol{\sigma}\rangle = \frac{1}{V}\int_\Omega \sigma \mathrm{d}V$

such that the effective tangent moduli can be expressed as:

$$\delta\boldsymbol{\sigma} = \bar{\boldsymbol{C}}^{\tan} : \delta\bar{\boldsymbol{\varepsilon}}\,, \quad \bar{\boldsymbol{C}}^{\tan} = \langle\boldsymbol{\mathcal{A}} : \boldsymbol{C}^{\tan}\rangle \tag{9}$$

considering $\langle\cdot\rangle$ as the volume average of the quantity $\cdot$ Note that the effective tangent modulus can be evaluated directly from the tangent stiffness matrix of the unit cell, such that the explicit evaluation of the localization tensors is not required. The computation of the tangent stiffness matrix is utilized to solve linear systems considering as boundary conditions separated perturbation of each component of the effective strain tensor, coupled with the nodes of each opposite face thanks to a set of Multiple Point Constraints, as presented for instance in [5].

In summary, the determination of the effective behavior of an heterogeneous material with periodic homogenization techniques relies mainly on two steps: a localization step where adapted boundary conditions (especially periodic conditions) are applied on the RVE and an homogenization step where the macroscopic response is obtained through volume averaging. These two steps repeated at each time step allows to establish the constitutive relationship between the macroscopic stress $\bar{\sigma}$ and the macroscopic strain $\bar{\varepsilon}$. This procedure will be detailed in the next subsection in the context of the multilevel finite element method FE^2.

2.2. Multilevel finite element method (FE^2)

To compute the effective non-linear response of heterogeneous structures while taking into account microstructures mechanical behavior, multilevel homogenization schemes such as the FE^2 method is commonly applied [2,3]. This approach, based on periodic homogenization theory, consists on solving the micro and macro problems simultaneously through localization and homogenization principles. The governing micro macro problem set of equations and scale coupling are summarized in Tables 1 and 2 respectively. As shown in Fig. 3, first an RVE is associated to each macrostructure Gauss integration point, then localization is performed to solve the periodic boundary value problem at the microscopic scale providing the RVE geometry and local constitutive equation, and finally the homogenization step is applied to obtain the macroscopic response. The main advantage of such approach lies in the fact that an explicit form of the macroscopic constitutive law is not required while solving the macro problem, the relationship between $\bar{\sigma}$ and $\bar{\varepsilon}$ is established solely through scale transition by computing the appropriate average quantities and the macroscopic tangent operators.

The resolution of non-linear multi-scale problems with FE^2 method is traditionally done in an incremental way using implicit resolution schemes such as Newton–Raphson at both scales. Fig. 4 illustrates the key steps to compute incrementally the macroscopic fields using the concept of periodic homogenization. At each time step (n), the local equilibrium problem is solved for every macroscopic Gauss integration point where an unit cell is associated. The macroscopic strain increment $\Delta\bar{\varepsilon}$ and periodic boundary

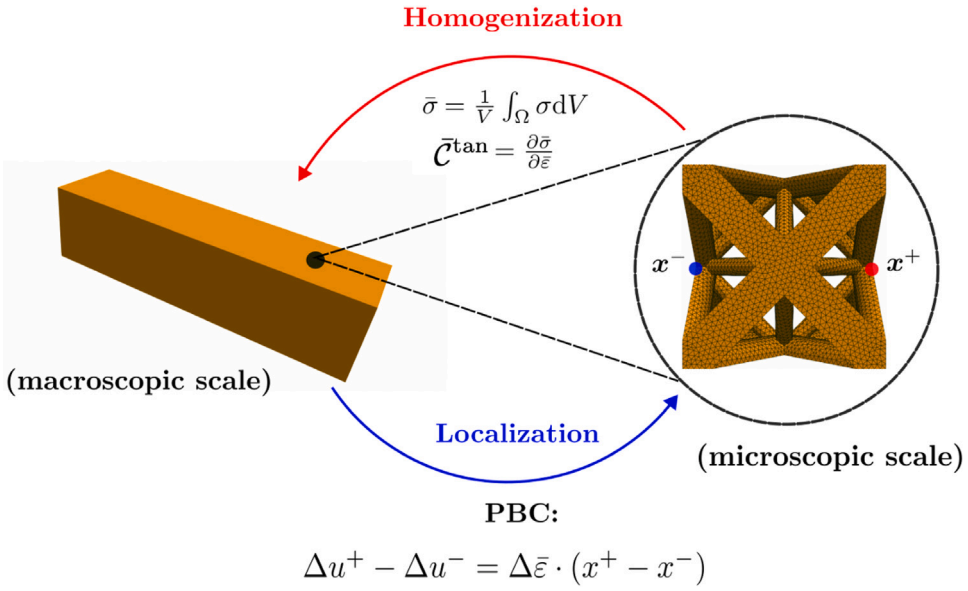

Fig. 3. Schematic representation of the micro-macro scale transition in FE^2 method.

conditions are applied to the RVE (localization step). Then, the BVP problem is iteratively solved using FE method and the microscopic quantities $\left(\varepsilon^{(\mathbf{n})}, \sigma^{(\mathbf{n})}, \zeta^{\mathbf{n}}\right)$ are updated for each microscopic Gauss point. $\left(\varepsilon^{(\mathbf{n})}, \sigma^{(\mathbf{n})}\right)$ are the micro strain and stress at the time step n respectively, $\zeta^{\mathbf{n}}$ is the set of internal state variables at the time step n. Finally, the macroscopic stress $\bar{\sigma}^{(\mathbf{n+1})}$ is computed through volume averaging of the microscopic stresses (homogenization step) and the incremental macroscopic tangent operator $\overline{\mathbf{C}}_{\mathbf{tan}}^{(\mathbf{n+1})}$ is computed using perturbation techniques. Once the macroscopic quantities ($\bar{\sigma}^{(\mathbf{n+1})}$, $\overline{\mathbf{C}}_{\mathbf{tan}}^{(\mathbf{n+1})}$) are calculated, the global equilibrium is checked. If the macroscopic convergence is not satisfied, a new macro strain increment $\Delta\bar{\varepsilon}'$ is provided by the FE solver and the local problem is solved once again. This process is performed until numerical convergence is achieved at both scales. In that case the multiscale analysis proceeds to the next time increment $(n+1)$.

Although the FE^2 can be considered as an attractive alternative compared to direct FE simulations on fully meshed heterogeneous structures, it is clear that this modeling strategy also suffers from some limitations given many factors. As described in the above mentioned FE^2 procedure, non-linear problems are solved for both the micro and macro scales using iterative schemes. This process is usually expensive due to the treatment of material non-linearities. In addition, other considerations may increase this computational time including the number of Gauss integration points, the complexity of loading conditions, the number of iterations and the meshing resolution of the RVE. Therefore, all these parameters may sometimes prevent this approach from being efficiently applied to structural analysis, which has led to the definition of optimized numerical implementations such as in [6]. This optimized technique allows a speed-up of about 60% which conserving all informations about the local fields. In the next section, we propose to develop a new numerical modeling strategy using a deep neural networks based method (FE-LSTM) in order to overcome the aforementioned restrictions.

3. FE-LSTM: A multi-scale approach combining the FE method and LSTM neural networks for heterogeneous structures modeling

3.1. An overview of long short term memory (LSTM) neural networks

Long Short Term Memory (LSTM) neural networks [36] are a class of Recurrent Neural Networks (RNN). Unlike Multi Layer Perceptrons (MLP), RNNs are designed to process sequences of data where the position of the data inside the sequence is important. This type of neural network architecture is useful for modern application such as Natural Language Processing (NLP), voice recognition, machine translation and time series analysis. LSTM were specifically designed to overcome vanishing and exploding gradient problems, which can occur when training traditional RNNs with Back Propagation Through Time (BPTT) algorithm [37]. The particularity of LSTMs compared to RNNs is its gating and adaptative memory mechanism that allows to control the flow of information. As presented in Fig. 5, a LSTM cell is composed of four main gates : An input gate i_t which controls the information added to the cell state, a forget gate f_t which controls how much information is kept or forgotten from the previous accumulated information, a memory gate C_t which represents the long term memory of the LSTM cell and the output gate o_t for the prediction of the next hidden state.

As shown in Fig. 5, the input vector to the LSTM unit at the time step t is x_t and the output vector is h_t. The parameters W_f, W_i, W_o, W_g, U_f, U_i, U_o and U_g are weight matrix corresponding to each gate, this set of parameters needs to be learned during the

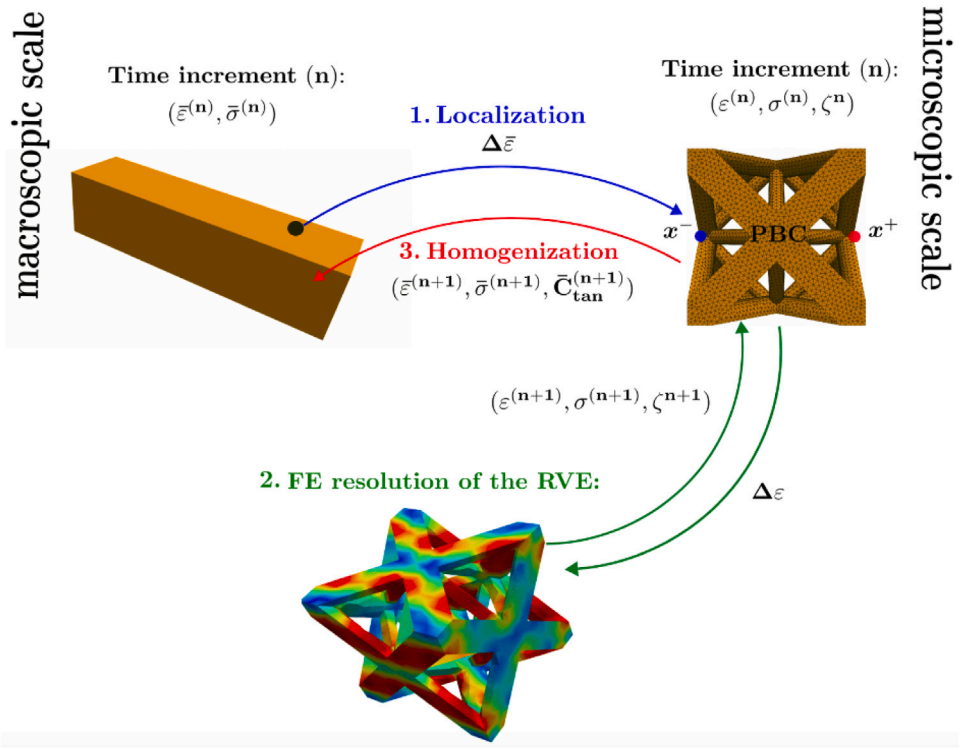

Fig. 4. Schematic representation of the FE^2 key steps while solving multi-scale problems incrementally.

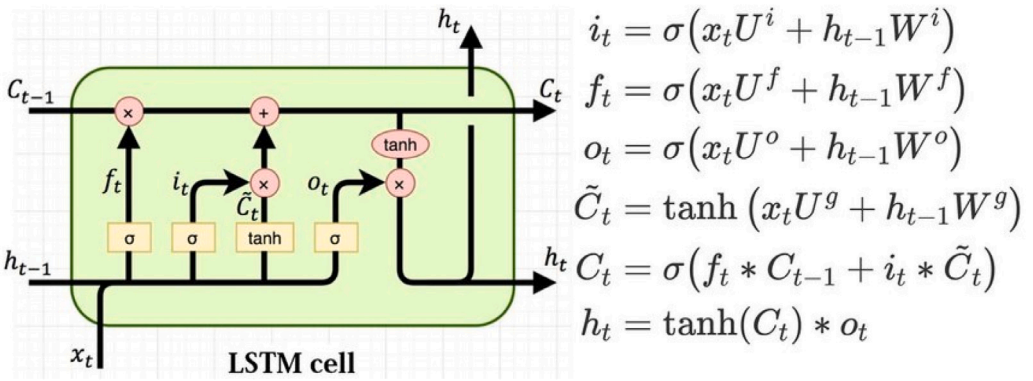

Fig. 5. Architecture of a LSTM cell.

backpropagation through time algorithm. The operator $*$ denotes the element-wise product, σ is the sigmoid function and tanh is the hyperbolic tangent function.

$$\sigma(x) = \frac{1}{1+e^{-x}},$$

$$\tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}.$$

3.2. FE-LSTM architecture design

The proposed approach, called FE-LSTM, combines the finite element method and an LSTM recurrent neural network to solve multi-scale problems. The principle of this methodology is to treat the microscopic and macroscopic problems separately: Unlike the FE^2 method, the FE resolution of the microscopic problems is no longer required and the computation of the RVE effective response is predicted by an LSTM trained on a database of offline micro computations, where the average macroscopic fields are post-processed.

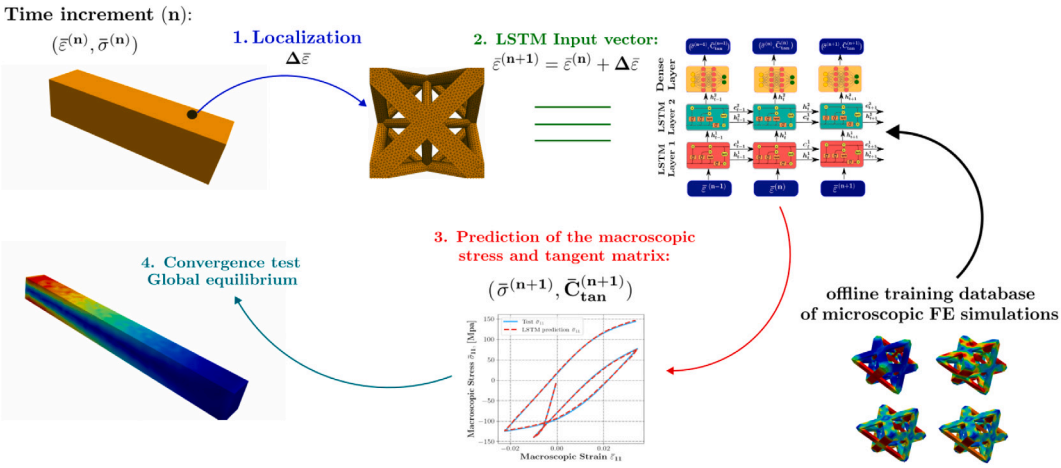

Fig. 6. FE-LSTM general workflow to compute the effective mechanical response of non-linear heterogeneous structures.

Using the predicted macroscopic fields, a finite element analysis is conducted to verify the global equilibrium at the macro-scale, this process is then performed until the achievement of numerical convergence. From the above description of FE-LSTM, it is obvious that the approach is heavily inspired by the FE^2 method as it mostly rely on the same workflow. However, the main advantage of FE-LSTM lies in the massive computational time savings when solving the microscopic problems. This performance can only be achieved with a reliable surrogate such as deep neural networks, especially considering non-linear models and non-proportional loadings. Based on the studies conducted in [24,32,33], the robustness of LSTM have been highlighted as a powerful tool capable of handling path dependent behaviors and accurately capturing complex and highly non-linear behaviors, hence the employment of such architecture to approximate the effective material responses in this multiscale framework.

It shall be noted though that without specific thermodynamic constraints, LSTM are in general not suitable to capture with a great precision the response of elastic–plastic materials during the first elastic stage that usually follow a change of direction in a non-proportional loading path [24]. In the proposed approach, since the constitutive model is a priori unknown, the enforcement of non-negative dissipation is not straightforward. However, in the framework of homogenization, a smooth response is generally observed on the macroscopic effective response, resulting from averaging operations. It is therefore expected that the response should be accurate enough to proceed with the simulation of the macroscopic response of structures.

The general workflow of the FE-LSTM approach is summarized in Fig. 6, this illustration describe the key steps to compute the macroscopic fields incrementally. First, an RVE is associated to all Gauss integration points. The macroscopic fields $\left(\bar{\varepsilon}^{(n)}, \bar{\sigma}^{(n)}\right)$ are known at the time increment (n), a macro strain increment $\Delta\bar{\varepsilon}$ is then applied to the unit cell. $\Delta\bar{\varepsilon}$ is added to $\bar{\varepsilon}^{(n)}$ to compute $\bar{\varepsilon}^{(n+1)}$ which is the input of the LSTM neural network. As shown in Fig. 7, the corresponding architecture is a Stacked LSTM composed of several LSTM layers and a dense layer at the end. This neural network takes $\bar{\varepsilon}^{(n+1)}$ as an input, then stores the material mechanical state in its memory gate, updates the input gate, forget gate, output gate, values to compute the next hidden state which is contains the effective stress and tangent operator $\left(\bar{\sigma}^{(n+1)}, \bar{\mathbf{C}}_{\mathbf{tan}}^{(n+1)}\right)$. Using these predicted quantities for all the Gauss integration points, the global equilibrium is checked using a FE analysis. Similarly to the FE^2 approach, if the convergence test is not successful, a new macro strain increment will be provided by the FE solver and new macroscopic fields will be predicted by the LSTM model. This process is performed until the satisfaction of numerical convergence and then the multiscale analysis proceeds to the next time increment. By default, the time increment and therefore the increment of values for the new boundary conditions is divided by four. A new time increment is performed with this value, making sure that all the field values are reinitialized to the beginning of the time increment. This operation is done by the solver fedoo automatically and no further operation is required. The only difference is that a new strain increment is taken locally by the LSTM, and there is a possibility that the FE-LSTM may not be accurate considering very small time increments since the exact solution is not found, and relative error can become significant in the case of such very small increments. Although this possibility, no convergence issues were observed during the numerical test cases investigated in Section 4.

4. Results and discussion

4.1. Architectured materials database generation

The first step within the FE-LSTM approach is the creation of a LSTM model. The calibration of the LSTM model parameters requires an offline training phase involving a database of numerical simulations on the RVE. Before proceeding to the description of the database generation strategy, some clarifications needs to be addressed regarding the main hypothesis of this study. First, the microstructure geometry is similar everywhere in the macrostructure, so that the same RVE is associated to all Gauss integration

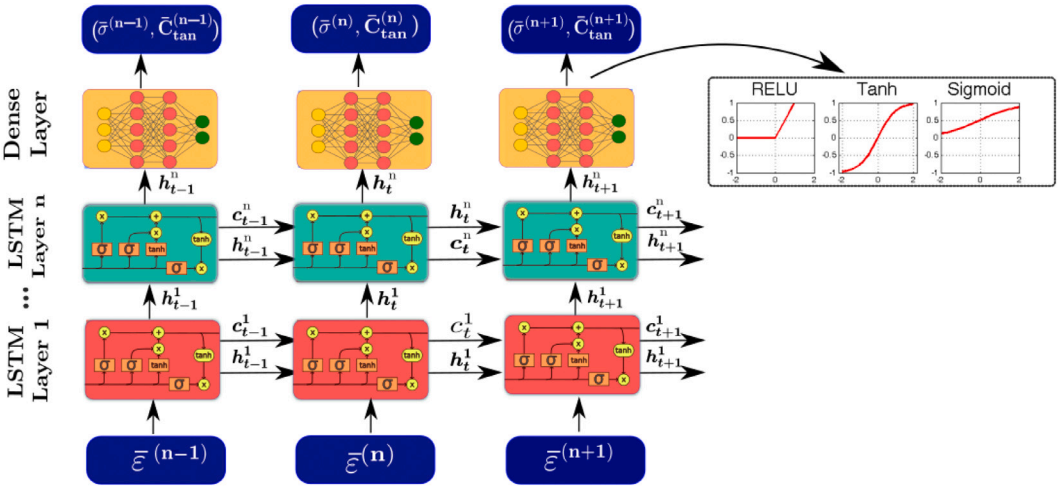

Fig. 7. Stacked LSTM Neural Network architecture to predict the homogenized response of the RVE.

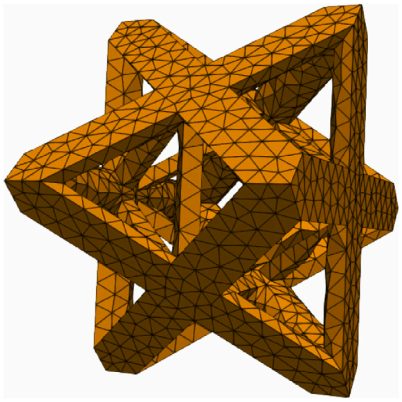

Fig. 8. Geometry and mesh of the octet truss RVE (Number of nodes: 1694, Number of elements: 8284).

Table 3
Mechanical properties of titanium alloy Ti–6Al–4V and geometrical properties of the considered octet truss.

Material parameter	Value
Young's modulus E	113 800 MPa
Poisson's ratio ν	0.34
Yield Stress σ_Y	1000 MPa
Hardening parameter H	1600 Mpa
Plastic hardening exponent n	0.5
External cylinders radius R_e	0.1
Internal cylinders radius R_i	0.05

points. The corresponding microstructure is an octet-truss architectured material whose geometry and periodic mesh are illustrated in Fig. 8. The geometrical and mechanical properties of the constitutive material which is a titanium alloy Ti–6Al–4V are summarized on Table 3. The non-linear microscopic constitutive equation is an elastoplastic law with isotropic hardening. Therefore, the only varying parameters while conducting RVE simulations is the applied loading conditions which are characterized by the evolution of the macroscopic strain tensors $\bar{\varepsilon}^{(n)}$.

The adopted strategy to generate the database $\overline{\mathscr{D}}$ is described as follows: 10 000 samples of microscopic RVE simulations, subjected to periodic boundary conditions and to multi-axial and non proportional loading paths, are generated using a finite element software *Fedoo* which relies on the *Simcoon* library to solve the constitutive equations [38]. The necessary resources to execute the simulations were provided by the computing facilities MCIA (Mésocentre de Calcul Intensif Aquitain) of Université de Bordeaux. The applied loading conditions on RVE boundaries are three-dimensional and can be expressed in terms of the six components of the macroscopic strain tensor $\bar{\varepsilon}^{(n)} = [\bar{\varepsilon}^{(n)}_{11}, \bar{\varepsilon}^{(n)}_{22}, \bar{\varepsilon}^{(n)}_{33}, \bar{\varepsilon}^{(n)}_{12}, \bar{\varepsilon}^{(n)}_{13}, \bar{\varepsilon}^{(n)}_{23}]$. To generate non-proportional loading paths, the total number of time

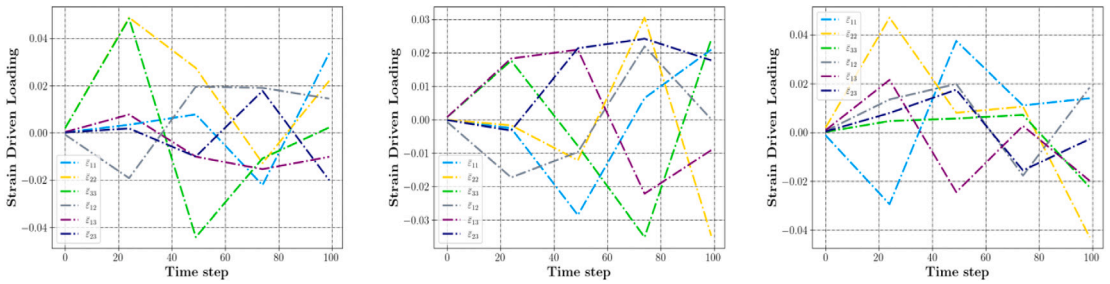

Fig. 9. Examples of multi-axial and non proportional loading conditions for RVE finite element simulations.

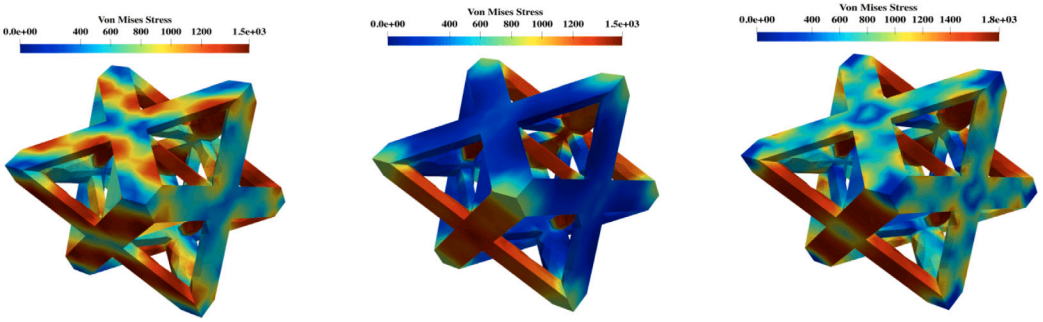

Fig. 10. Examples of simulated octet-truss using FE analysis, the corresponding homogenized fields are used as training data.

Table 4
Evaluation of LSTM configurations training and validation performances for the macroscopic stress–strain responses prediction.

ANN configuration	Number of hidden LSTM layers	Output layer activation function	Training loss NMSE (%)	Validation loss NMSE (%)
Config 1	2	Relu	0.011%	0.009%
Config 2	3	Relu	0.007%	0.006%
Config 3	2	Sigmoid	0.013%	0.011%
Config 4	3	Sigmoid	0.009%	0.008%
Config 5	2	Hyperbolic tangent	0.013%	0.013%
Config 6	3	Hyperbolic tangent	0.011%	0.010%

steps, which is fixed at 100, is divided into 4 steps of 25 increments. Each step is linear and the bounds values of each macroscopic strain are chosen randomly from the interval $[-5\%, 5\%]$. Fig. 9 shows some examples of 3D loading path configurations used as training data. Some results of unit cells finite element computations are illustrated in Fig. 10: the associated microscopic fields are subsequently homogenized to obtain the macro fields. Finally, the database $\overline{\mathscr{D}}$ is constructed using the set of macroscopic quantities $(\bar{\varepsilon}, \bar{\sigma}, \overline{\mathbf{C}}_{\mathbf{tan}})$ which are fed as time sequences to train the LSTM model. As usually, $\overline{\mathscr{D}}$ is split into a training set $\overline{\mathscr{T}}$ (90% of total samples), a validation set $\overline{\mathscr{V}}$ (20% of $\overline{\mathscr{T}}$) and a test set $\overline{\mathcal{T}}$ (10% of total samples).

4.2. Model training phase and hyperparameters selection

Following the database generation and features scaling, the next step is the training process of LSTM models. For this purpose, two LSTM models have been created to predict separately each quantity of interest, i.e the macroscopic stress $\bar{\sigma}$ and the macroscopic tangent stiffness $\overline{\mathbf{C}}_{\mathbf{tan}}$. This approach is much more practical than creating the same model to predict both quantities as they are of different nature, and thus capturing the non-linear behaviors may require different parameter adjustments dependant on the complexity of each model. Models implementation was carried out using *Keras* library [39] and Tensorflow API [40]. Six different configurations of LSTM architectures were tested by varying three activation functions (ReLU, Hyperbolic tangent and Sigmoid) and the number of LSTM layers. This hyper-parameter study aims to identify the most reliable model that minimizes training and validation errors. The common Mean Squared Error (MSE) was employed as a loss function and Adam (optimizer with adaptative moment) was used as an optimization algorithm.

For the prediction of macroscopic stress tensors, the evolution of training and validation Normalized Mean Squared Error (NMSE) metric are shown in Fig. 11 for all the six tested configurations. It can be noticed from the corresponding graphs that the learning process is achieved rapidly regardless of the choice of the activation or the number of LSTM layers. A significant decrease in the errors is observed during the first epochs, followed by a loss stabilization after nearly 20 epochs. The minimum values of training and

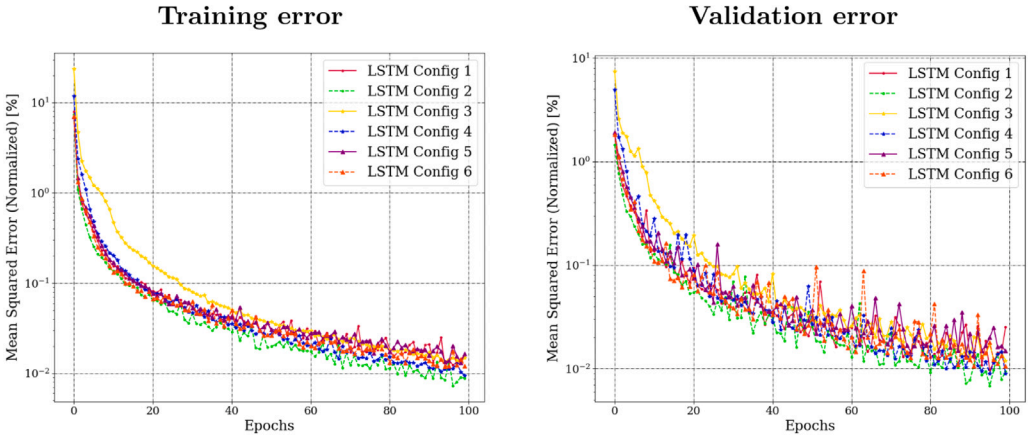

Fig. 11. Evolution of training (left figure) and validation errors (right figure) for the tested LSTM configurations (Prediction of the macroscopic stress).

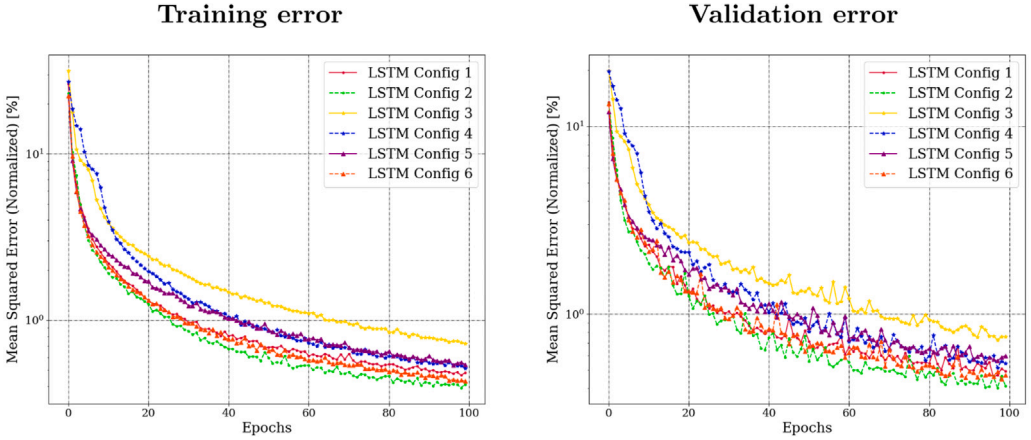

Fig. 12. Evolution of training (left figure) and validation errors (right figure) for the tested LSTM configurations (Prediction of the macroscopic tangent matrix).

Table 5
Evaluation of LSTM configurations training and validation performances for the tangent matrix predictions.

ANN configuration	Number of hidden LSTM layers	Output layer activation function	Training loss NMSE (%)	Validation loss NMSE (%)
Config 1	2	Relu	0.41%	0.42%
Config 2	3	Relu	0.39%	0.40%
Config 3	2	Sigmoid	0.72%	0.72%
Config 4	3	Sigmoid	0.51%	0.51%
Config 5	2	Hyperbolic tangent	0.53%	0.55%
Config 6	3	Hyperbolic tangent	0.42%	0.44%

validation loss obtained through all the learning process are summarized in Table 4. We can observe that the error values are almost the same for all configurations (order of magnitude of 10^{-2}%). Furthermore, these low error values are an evidence that the model has achieved excellent training performances. Even though different activation functions and number of LSTM layers can be used to approximate the macroscopic stress–strain response, the first configuration has been chosen to fix model hyper-parameters. This choice is justified by the following reasons: First, two LSTM layers instead of three are sufficient to obtain a good approximation. An increase of number of layers does not necessarily improve model performance. In addition, the prediction of the quantities of interest is much faster using a small number of layers, which is consistent with the objectives of the study. For the activation function selection, the Rectified Linear Unit ReLU is computationally more efficient in back-propagation compared to Sigmoid and Hyperbolic tangent functions as it only takes the maximum between the input and zero value. The required training time with ReLU is nearly three times faster than training with Sigmoid or Hyperbolic tangent. Furthermore, ReLU activation is more suitable for overcoming vanishing gradient problems, thus allowing LSTM models to learn faster and perform better (see Fig. 12 and Table 5).

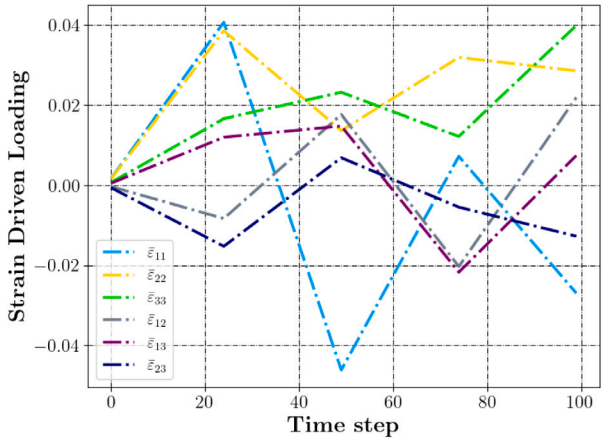

Fig. 13. The applied multi-axial and non proportional loading path on RVE boundaries for an example of test data.

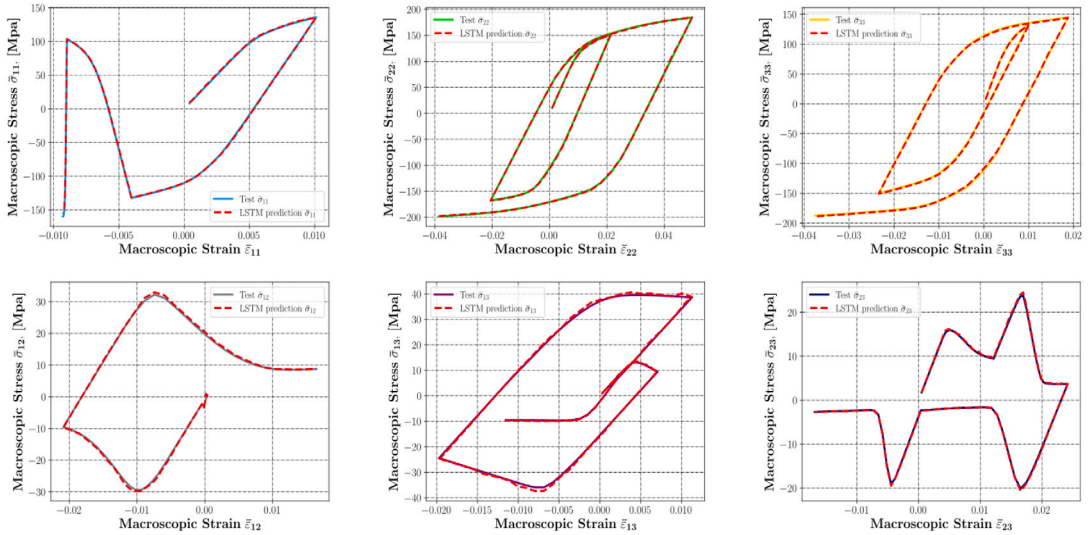

Fig. 14. LSTM predicted macroscopic stress–strain response in comparison to the RVE FE simulation for an example of test data.

Regarding the training process on macroscopic tangent operators, the same observations apply to this model about the loss evolution and the errors order of magnitude being almost the same for all configurations. However, it can be noticed that the resulting errors are much larger compared to the values obtained with the previous model. There are two main differences between these models : (i) the number of the predicted outputs (21 components of the symmetric part of the tangent matrix) is much higher than the number of macroscopic stress components. When the number of outputs is large, the model is more expressive, and may describe statistical peculiarities of the training data that are atypical and lead to unusual predictions, which is known as *overfitting* [41]; (ii) The level of complexity and non-linearity of the response is different. Finally, the numerical noise resulting from the computation of tangent operators (especially for the insignificant components) can also be a cause of error amplification. However, it is important to note that training and validation error values regarding this quantity of interest are still very low (0.4% for the first configuration), thus not having a significant effect on the model's reliability.

4.3. Prediction of the macroscopic mechanical behavior of octet-truss structures

In order to evaluate the performance and the generalization capabilities of LSTM models, it is crucial to verify their predictive accuracy on the test dataset $\overline{\mathcal{T}}$. We recall that this batch of data (10% of total samples) was never employed during the training phase, hence it is a good indicator to determine if the model suffers from potential overfitting. Fig. 14 presents an example of the LSTM predicted macroscopic stress–strain response in comparison to the corresponding RVE finite element solution under a multi-axial and non proportional loading path shown in Fig. 13. The obtained results in this example shows an excellent agreement between the predicted macroscopic stress components and the target values. The model assess a strong capability to capture the non-proportional,

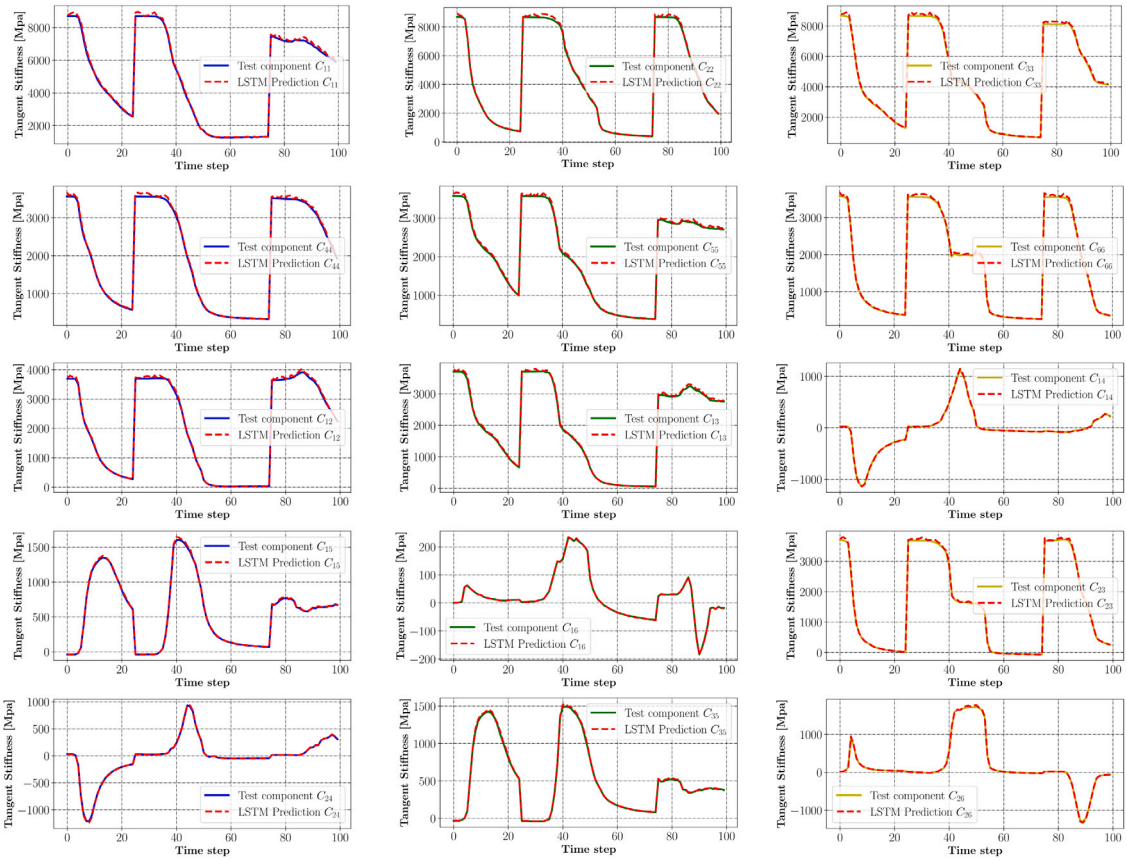

Fig. 15. LSTM predicted macroscopic tangent matrix components in comparison to the RVE FE simulation for an example of test data (Only the symmetric part and the significant components are showed).

non-linear responses despite the complexity of the loading condition. In addition, it can be seen from Fig. 14 graphs a complete absence of numerical noise during a change of direction in the stress–strain curves, which is advantageous for the convergence of multiscale simulations. To quantify the model's accuracy on the entire test data samples, we use the Mean Absolute Percentage Error (MAPE) metric expressed by Eq. (10). The resulting MAPE for the prediction of macroscopic stress–strain responses is found to be equal to 0.27%. As a consequence the LSTM model achieved a very high accuracy rate ($\approx 99\%$). This finding emphasizes once again the potential of RNN-based models as reliable surrogates capable of capturing history dependant non-linear behaviors. This preliminary outcome is very encouraging for the next step of this study where LSTM shall be integrated in a multiscale approach within the FE-LSTM framework.

$$MAPE = \frac{1}{N}\sum_{k=1}^{N}\frac{\left|y_{\text{true}}^{(k)} - y_{p}^{(k)}\right|}{|y_{\text{true}}^{(k)}|} \times 100 \tag{10}$$

- $\boldsymbol{y}_{true}^{(k)}$: Target values
- $\boldsymbol{y}_{p}^{(k)}$: Predicted values
- N : Total number of test samples

We also assess the reliability of the second model for the prediction of macroscopic tangent matrix $\overline{\mathbf{C}}_{\mathbf{tan}}$. A good evaluation of the tangent operators is critical for the computation of non-linear heterogeneous structures within Newton–Raphson framework. Fig. 15 illustrates the comparison between tangent modulus components predictions and the target values from simulation for an example of test data. We can observe from this example a good correspondence between the two quantities, the general trend of the tangent moduli components is well captured by the LSTM during all the time steps. However, the level of accuracy obtained for this quantity of interest is not as high as the precision achieved for the stress components, which is expected given the number of outputs. In addition, it can be noticed from Fig. 15 graphs a presence of small noise especially in the first time increments Note that while it is difficult to provide a clear explanation given the complexity of the training process, LSTM network still do not completely prevent against the vanishing gradient phenomena, which appears in the first stage of a sequence of data, given the recursive data analysis process. Nevertheless, it is important to note that these errors do not have a large influence on the overall model reliability, the evidence is that the obtained MAPE on the test data samples is equal to 2.57%, which is still a low value.

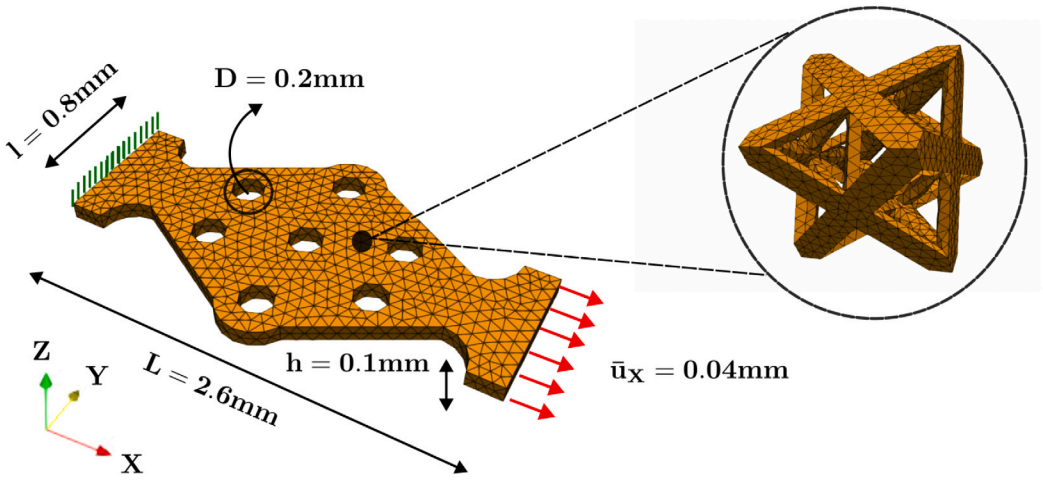

Fig. 16. Geometry, mesh and boundary conditions of an architectured specimen with holes under tensile loading.

In the next step of this study, we propose to apply this approach for the structural analysis of 3D architected structures considering engineering Boundary Value Problems. Since the RVE loading states are a priori unknown during multi-scale simulations, the challenge that arises is whether LSTM models have the ability to capture with a sufficient accuracy the macroscopic behavior when the RVE are subjected to loading paths that are not necessary based on linear steps employed in the training process.

4.4. Application to a complex architectured structure under proportional and non proportional loading paths

4.4.1. Case of a proportional loading path

To illustrate the flexibility of the FE-LSTM approach when applied to complex structures, two application cases of a non-standard architectured specimens under proportional and non-proportional loading are tested in this subsection. The first example consists of a 3D specimen in the form of a plate drilled in several locations to enhance a non-uniformity of the mechanical field, subjected to a uniaxial tensile loading. The geometry is inspired from [42], which has been used to validate a numerical homogenization scheme using ROM and ANN. The macrostructure geometry, mesh and boundary conditions are presented in Fig. 16. Dirichlet boundary conditions are applied to the heterogeneous structure in the x direction. One extremity is clamped while the other is subjected to a displacement-controlled loading. All other faces are traction-free. The specimen is meshed using 4-node tetrahedral elements TET4 with 4 Gauss integration points per element and the full mesh is composed of 13 400 Gauss points. The computation of the macroscopic mechanical response of the heterogeneous structure is performed with FE-LSTM model and using FE^2 method. Both approaches are compared according to different criteria including the accuracy of the mechanical fields, simulation time and the required memory usage. Since the FE^2 method is very time consuming, the corresponding simulations are conducted using only 10 time increments. The analysis results are presented in Figs. 17 and 18. Fig. 17 illustrates the distribution of the macroscopic Von Mises stress $\bar{\sigma}_{VM}$, normal stress $\bar{\sigma}_{XX}$, $\bar{\sigma}_{YY}$ and shear stress $\bar{\sigma}_{XY}$ obtained by FE^2 and FE-LSTM. Regarding the macroscopic strain fields, Fig. 18 shows the distribution of the normal strain components $(\bar{\varepsilon}_{XX}, \bar{\varepsilon}_{YY})$ and shear strain component $\bar{\varepsilon}_{XY}$ respectively obtained by FE^2 and by FE-LSTM.

As it can be seen from both figures, an excellent correspondence is observed between the macroscopic fields predicted by FE-LSTM model and the targets computed by FE^2. Despite the complexity of the geometry, the developed model is able to accurately predict the non-linear homogenized response of the architectured structure. Similarly to the FE^2 approach, FE-LSTM is also capable of taking into account the effect of the microstructure on the overall mechanical response of the macrostructure. In addition, we can notice from the distribution of the macroscopic stress field in the loading direction $\bar{\sigma}_{XX}$ a stress concentration effect due to the presence of holes. This stress localization is also captured by the FE-LSTM in the specimen reduced cross-sectional area and its neighboring holes, which is expected given the applied boundary conditions. In terms of values accuracy of the predicted macroscopic fields, a very good agreement between the two approaches can be observed as well, given the low error values (MAPE) reported in Table 6. Furthermore, the maximum deviation of the macroscopic stress $\bar{\sigma}_{XX}$ between FE-LSTM and FE^2 for a Gauss point is at worst equal to 1%, which remains a reasonable error. From the distribution of the macroscopic strain fields shown in Fig. 18, we can verify that the strain components remain in the range of training data i.e in the interval $[-5\%, 5\%]$. As neural networks are known to have poor extrapolation capabilities, it is always important to check this condition since a deformation state outside the training database could result in inaccurate predictions or a failure in convergence of the simulation.

We recall that the main objective of this study is the acceleration of multi-scale simulations, therefore a breakdown of the offline and online computational costs is required to highlight the efficiency of the developed approach FE-LSTM.

The database generation was carried out using the computing facilities of the cluster MCIA, 10 nodes of 32 CPU each were involved in the creation of the database. The total resulting computational time (CPU time) is 70 h. The CPU time is expressed as

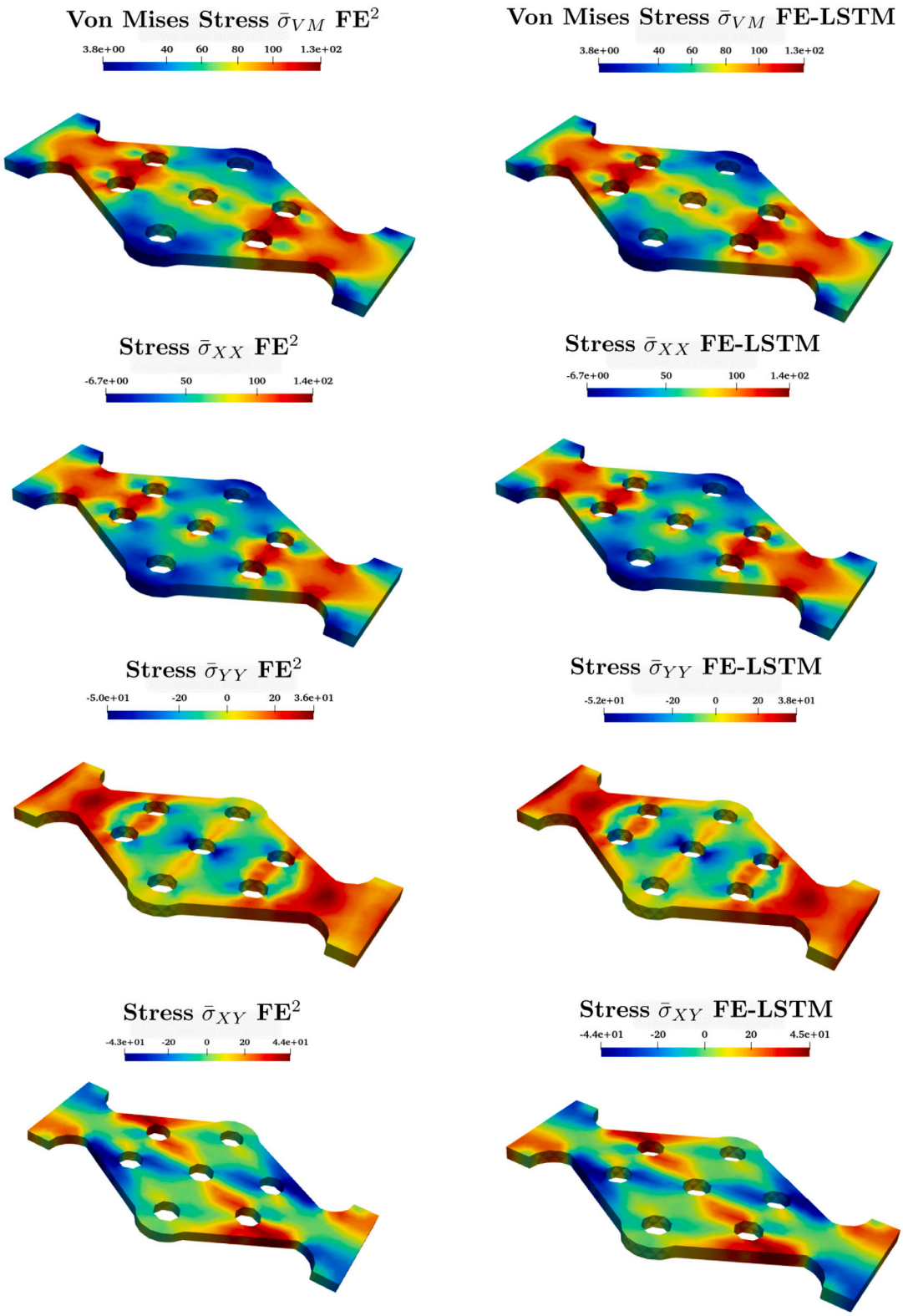

Fig. 17. Comparison between the macroscopic stress fields $\bar{\sigma}$ obtained by FE^2 and by FE-LSTM of an architectured specimen with holes under tensile loading.

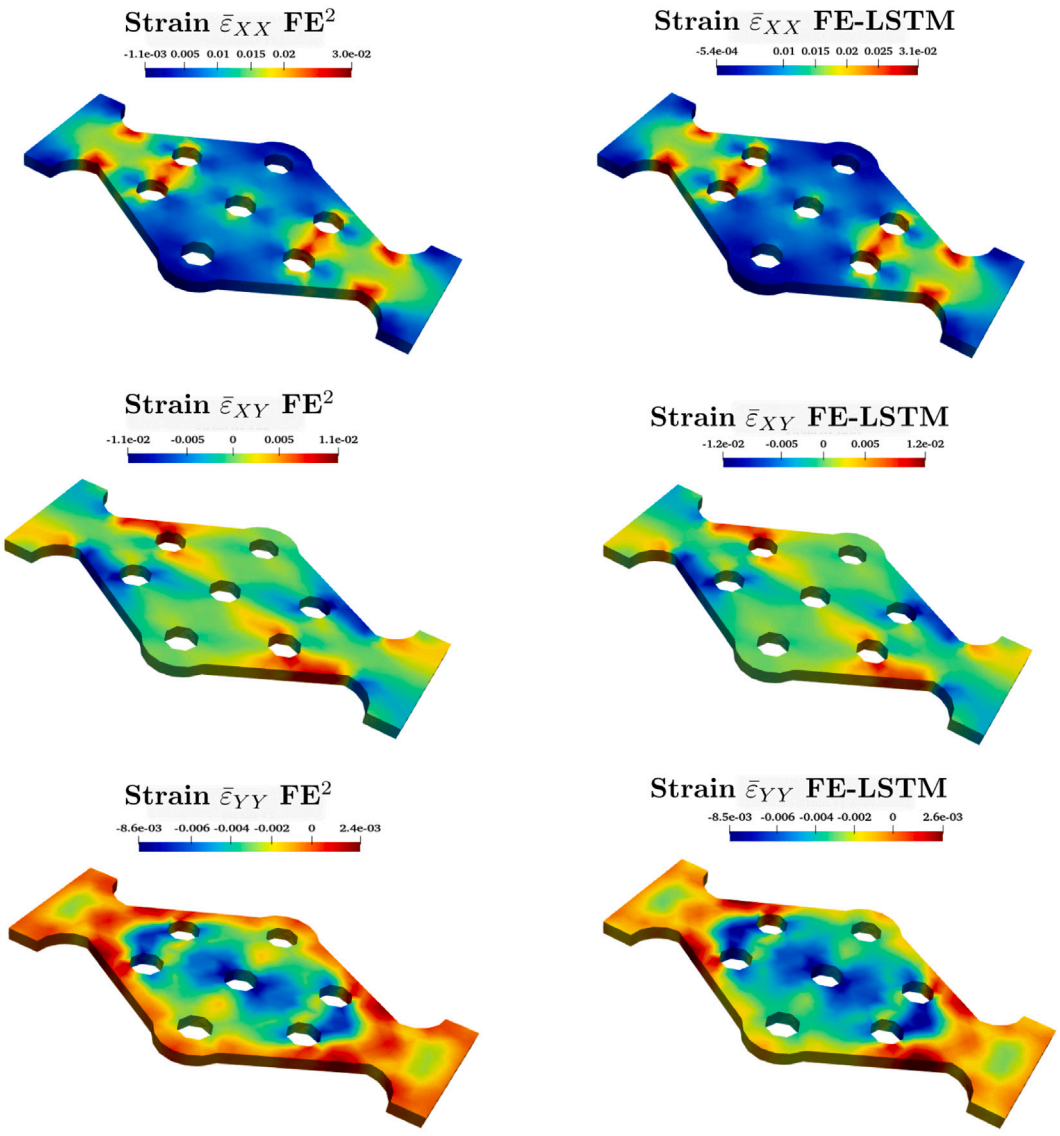

Fig. 18. Comparison between the macroscopic strain fields $\bar{\varepsilon}$ obtained by FE^2 and by FE-LSTM of an architectured specimen with holes under tensile loading.

Table 6
Macroscopic stress components MAPE values.

Macroscopic stress components	Stress $\bar{\sigma}_{VM}$	Stress $\bar{\sigma}_{XX}$	Stress $\bar{\sigma}_{YY}$	Stress $\bar{\sigma}_{XY}$
MAPE (%)	0.03%	0.03%	1.58%	1.97%

a sum of the CPU resources considering all simulations carried. Although this offline procedure remains expensive, major speed up can be achieved during the online computing stage. In addition, once the model is trained, it can be applied to any heterogeneous structure having the same microstructure. In contrast, the FE^2 method require solving once again non-linear problems on each macroscopic integration point at each iteration for any new simulation. FE-LSTM based approaches are therefore more interesting in the design of heterogeneous structures for examples, as they allow to quickly perform parametric analysis by varying the structure geometry or the loading conditions with respect to the desired performances.

In Table 8, the total computational time, that includes database generation computing time, LSTM training time and the online computational time is compared with the legacy FE^2 implementation. It is noted that the total time is still lower in this case for the FE-LSTM approach, with a speed up factor of 1.7. In this specific case, it is less than the monolithic implementation of FE^2 proposed by [6]. Considering only a single multi-scale simulation, it is therefore not relevant to consider the proposed framework, since a monolithic implementation of the FE^2 method is faster and provide also the whole information about the local fields, which are not

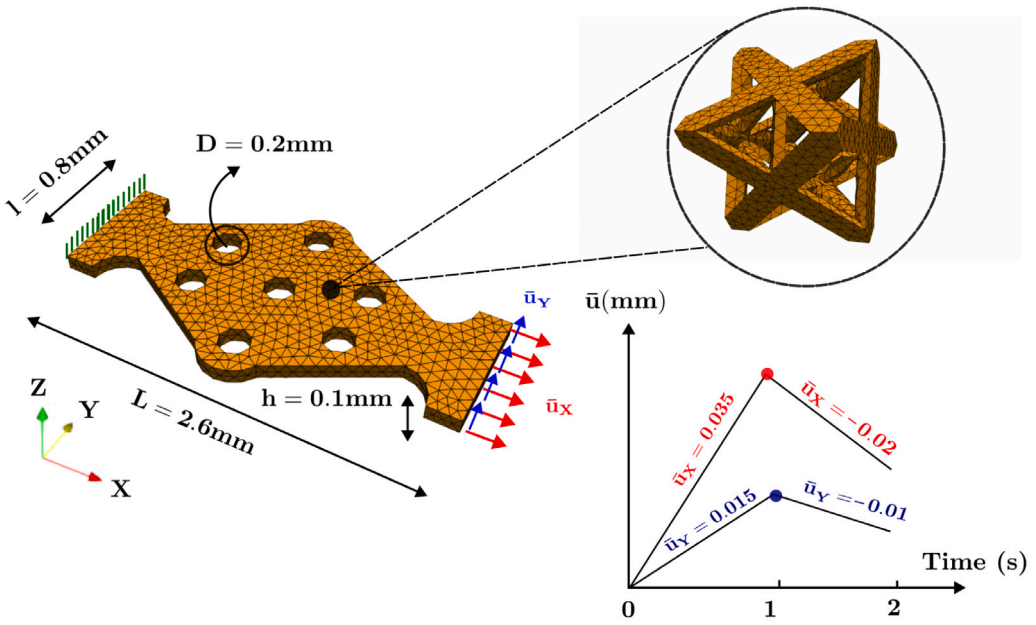

Fig. 19. Geometry, mesh and boundary conditions of an architectured specimen with holes under tensile and shear loading/unloading conditions.

Table 7
The required computational time and memory usage to simulate the non-linear response of the architectured structure using FE^2 and FE-LSTM.

	FE^2 simulation	FE-LSTM simulation	Computational time saving factor	Memory usage saving factor
Online simulation time	120 h	10,2 s = 0.0028 h	42 350	–
Memory usage	1.07 TB	120 MB	–	8917

accessible using the FE-LSTM approach. However, for the online computing stage, Table 7 summarizes the required computational time and memory usage for both approaches FE-LSTM and FE^2. The simulation of the heterogeneous structure took nearly 5 days using FE^2 while it only took 10.2 s using FE-LSTM, thus resulting in a computational time saving factor of 42 350. This speed up factor is higher than the latest implementation of the monolithic solution algorithm combine with reduced order modeling (ROM) and the empirical cubature method (ECM) [8]. If the online computational time is considered as the reference time (i.e. for topology optimization, parametric analysis, multiple loading analysis, decision-making assistant digital twin), the proposed solution is attractive. Furthermore, note that the achievable gains with FE-LSTM are not only related to the gain in computational time, but also to the computing resources. As shown in Table 7, FE^2 simulation required a memory usage of 1.07 TB while the FE-LSTM only used 120 MB. The FE^2 simulation involved the use of a *bigmem* machine of 3TB RAM from the MCIA cluster, which is not always affordable. In contrast, the FE-LSTM simulation could be carried out in a desktop computer without any need for expensive computing resources. The selection of the most appropriate algorithm strongly depends on the application, considering the number of simulations, the possible change (or not) of the microstructure, the size of the micro cell and its required accuracy, the size of the macroscopic structure and its mesh, the requirements in terms of accessible CPU time and RAM memory. The proposed solution appears as a interesting solution when only the macroscopic information is relevant, with an important number of simulations considering the same microstructure. In this case, a massive time saving can be observed, avoiding the necessity to utilize costly computing supercomputers to reach the RAM memory requirements.

4.4.2. *Case of a non-proportional loading path*

Within this second case of application, we aim to evaluate the reliability of the FE-LSTM approach for heterogeneous structures subjected to non-proportional loading. The same non-standard architectured specimen with several holes is considered as before, the only change is the applied loading conditions. The macrostructure geometry, mesh and boundary conditions are presented in Fig. 19. Dirichlet boundary conditions are applied to the heterogeneous structure in the direction x, one extremity is clamped while the other is subjected to a displacement controlled loading involving tensile and shear. All other faces are traction-free. This non-proportional path is composed of a tensile and a shear loading step with a final displacement value of $\bar{u}_X = 0.035$ mm and $\bar{u}_Y = 0.015$ mm respectively, followed by an unloading phase with displacement of $\bar{u}_X = 0.02$ mm and $\bar{u}_Y = 0.01$ mm. The simulation of the macroscopic response of the architectured structure is only performed using FE-LSTM. The FE^2 simulation was not conducted due to computational resources limitations, and would have required 10 days of simulation on a *bigmem* CPU of the MCIA cluster

Table 8
Total, offline and online computational costs using FE^2 and FE-LSTM.

	Total	Database generation	LSTM training	Online
FE^2	120 h	–	–	120 h
FE-LSTM	75.5 h	70 h	5.5 h	0.0028 h

for a total of 20 time increments (10 for the loading step and 10 for the unloading step). However, we would like to remind that the FE-LSTM approach has already been validated through two cases of application, and the case presented here is to show the predictive capabilities of our approach to simulate multi-scale engineering problems.

The resulting macroscopic Von Mises stress $\bar{\sigma}_{VM}$, and stress components $\bar{\sigma}_{XX}$, $\bar{\sigma}_{YY}$ and $\bar{\sigma}_{XY}$ obtained with FE-LSTM are presented in Fig. 20. The stress fields distribution on the architectured specimen are given at the end of tensile/shear loading and unloading steps. In absence of a FE^2 reference solution, we evaluate the numerical results in a qualitative way by verifying different aspects. First, the model achieved numerical convergence with the same Newton–Raphson tolerance as used in the previous applications. This first outcome indicate that the predicted macroscopic tangent operators are potentially well predicted by LSTM model despite the complexity of loading conditions (non-proportional) compared to a simple uniaxial tensile test (proportional). Due to the effect of tension, a stress concentration is observed in the specimen reduced cross-sectional area and in its neighboring holes. The effect of combining shear and tension is also captured by the model given the high stress state in zone A (see Fig. 20). This reduced area is subject to a maximum stress as expected considering the specimen geometry and loading conditions. The stress field are also continuous which indicate that no specific local stress artifact is observed in isolated elements. In the unloading phase, a stress concentration is also observed around the central hole (Zone B in Fig. 20) with a rather particular distribution due to the shear unloading effect. Based on the present findings, we can assume that FE-LSTM can also be adapted for solving history-dependent multi-scale problems, provided that the loading states remain in the LSTM training range.

4.5. Illustration of the achievable capabilities with FE-LSTM model for highly refined meshed architectured structures

The last example of the present work is intended to illustrate the achievable performance by FE-LSTM approaches for heterogeneous structures with highly refined mesh. This investigation is motivated by the following reason: Refined meshes (sometimes involving millions of degrees of freedom) and small increments are sometimes required to ensure a numerical convergence for structures with complex geometry and boundary conditions. The expected computational time can be considered as incompatible with respect to the engineering design timeframe. In addition, this kind of simulations are very demanding in terms of computing resources (memory, number of CPUs). It requires sometimes the usage of high performance machines with characteristics that are not easily accessible. We propose here, through this last engineering example, to show the possibility of conducting this kind of multiscale simulations without the use of dedicated numerical resources, i.e only with a desktop computer using the FE-LSTM approach. The example considered in this study is a bracket whose geometry, associated RVE, mesh and boundary conditions are given in Fig. 21. The bracket is meshed using 4-node tetrahedral elements TET4 with 4 Gauss integration points per element. The macrostructure mesh is composed of 282 824 Gauss points. The heterogeneous structure is clamped on one extremity (Face B) and subjected to a compressive loading at the other end (Face A). In Fig. 22, the Von Mises stress $\bar{\sigma}_{VM}$ distribution on the architectured bracket is given. Qualitatively speaking, the obtained results seem very accurate and consistent given the applied loading and boundary conditions. The stress concentration around the central hole is well captured by FE-LSTM model and the distribution of the stress fields is rather smooth.

As shown in Table 9, the execution of this simulation with FE-LSTM model took 3.5 min (using 10 time increments). On the other hand, the same simulation would have taken 105 days using a classic FE^2 thus resulting in a speedup of 43 200. This factor is calculated based on the required computational time of FE^2 in the previous example (architectured specimen with holes) extrapolated to this case based on the number of DOF of this example. Note that this factor assumes that FE^2 was not optimized for parallel computing and without specific treatment. Another interesting feature of FE-LSTM approach is that the total simulation time including database generation, training phase (5 h) and online simulation (3.5 min) is still insignificant compared to required FE^2 computational time (105 days). In this kind of application that involves highly refined meshes, there exists a threshold at which it is more interesting to train an LSTM model using a limited number of microscopic evaluations than to perform the entire multiscale simulation using FE^2, even for a single Boundary Value Problem to solve. In terms of memory usage, the required RAM to compute the homogenized response is estimated 22.58 TB with FE^2 while FE-LSTM only need 2.5 GB, which is compatible with typical desktop computers characteristics. 22 TB of RAM makes the simulation challenging to carry out as even *bigmem* machines are usually limited to 3 TB of RAM. Conducting this type of multiscale simulation requires specific treatment of the memory and usually involve domain decomposition methods. To conclude this subsection, we highlighted through an illustrative example the flexibility of FE-LSTM approach for applications involving highly refined (sometimes exaggerated) meshes. Even for these extreme cases, the FE-LSTM method is not affected by a significant increase in computational time: the multiscale simulations are still executed within minutes. In addition, they can be easily conducted on desktop computers without the need for clusters or high performance machines.

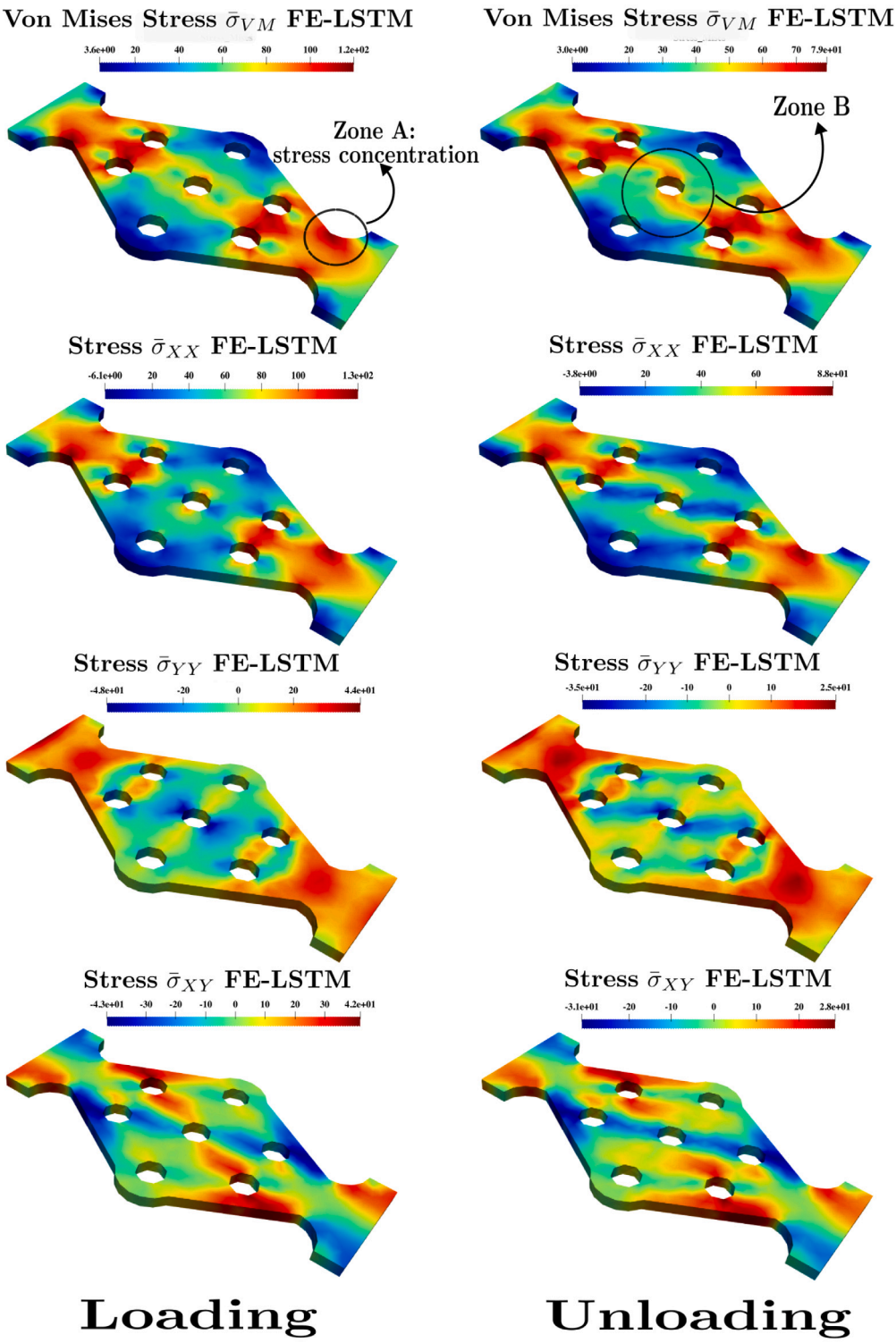

Fig. 20. The distribution of macroscopic stress fields $\bar{\sigma}$ obtained by FE-LSTM at the end of the loading and unloading steps.

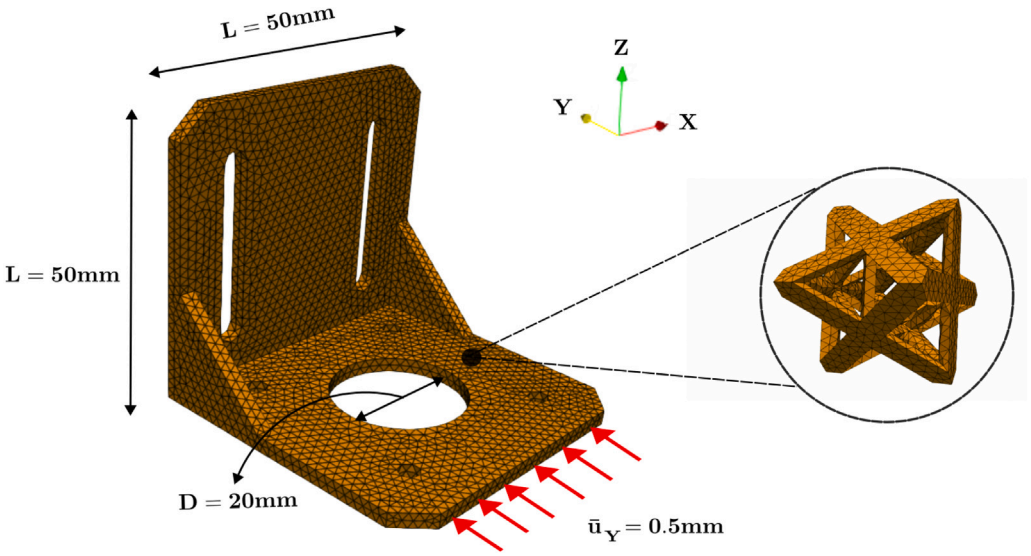

Fig. 21. Geometry, mesh and boundary conditions of an architectured bracket under a compressive load.

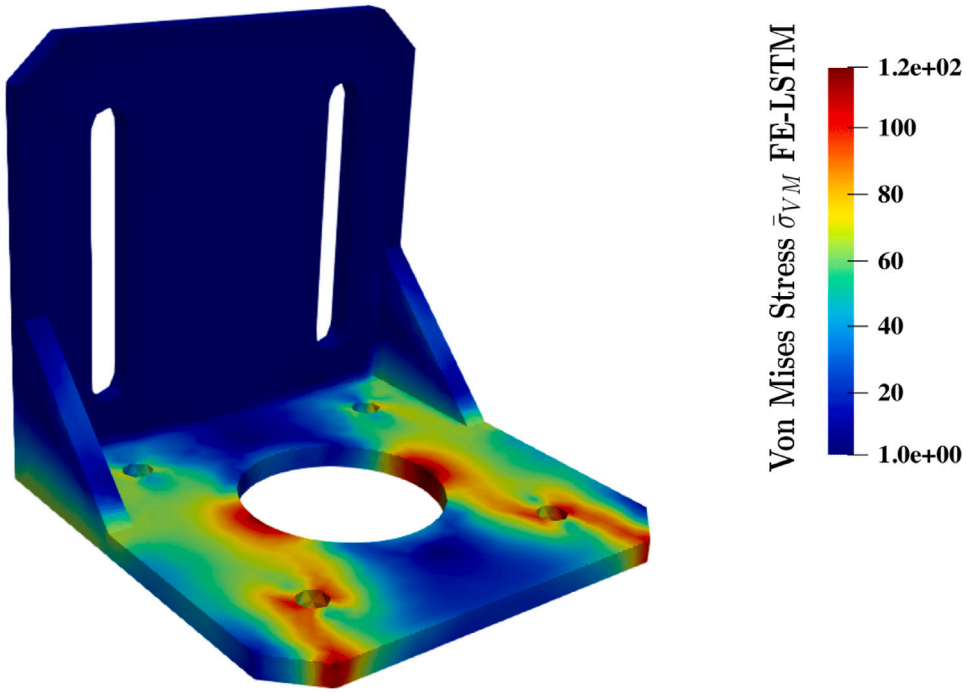

Fig. 22. Macroscopic Von Misses stress $\bar{\sigma}_{VM}$ distribution.

Table 9
The required computational time and memory usage to simulate the non-linear response of an architectured bracket structure using FE^2 and FE-LSTM.

	FE^2 simulation	FE-LSTM simulation	Computational time saving factor	Memory usage saving factor
Required simulation time	105 days = 2520 h	3.5 min = 0.58 h	43 200	–
Memory usage	22.58 TB	2.5 GB	–	9032

5. Conclusions and perspectives

Within this paper, we elaborated a novel modeling strategy to accelerate multi-scale simulations of architectured materials. The developed approach, called FE-LSTM, consists on combining the finite element method and LSTM recurrent neural networks to solve non-linear multi-scale problems. In contrast to FE^2 method, the FE resolution of the microscopic problems is no longer required within FE-LSTM framework, the computation of RVE homogenized response is directly predicted by an LSTM trained on a database of offline micro simulations. Consequently, the main advantage of FE-LSTM lies in the massive computational time and resources saving when solving the microscopic problems. However, the proposed approach predicts only the average macroscopic field, which is a limitation for instance if fatigue analysis, based on the heterogeneous local field at the microscopic scale, is required.

To assess the validity and reliability of the developed approach, FE-LSTM has been evaluated on 3D architectured structures under proportional and non proportional loading conditions. The obtained results have demonstrated FE-LSTM high predictive capabilities, an excellent agreement has been found between the macroscopic fields predicted by FE-LSTM and the targets computed by FE^2. The usefulness of such approach lies in the fact that once the RNN training process is performed, FE-LSTM can be applied to simulate nonlinear behaviors of any heterogeneous structure having the same microstructure used in training. By performing parametric analysis while testing a variety of macroscopic geometries and loading conditions, such approach find its interest in the design of heterogeneous structures as it allows to conduct multiple multi-scale simulations in a short amount of time. In terms of execution time, we have highlighted through the previous applications that FE-LSTM simulation are performed within seconds compared to days with FE^2, thus resulting in computational time saving factors of nearly **40 000**. Furthermore, they can be easily conducted on desktop computers without the need for HPC clusters.

Even though FE-LSTM achieved very promising performances based only on RNN high predictive capabilities and abilities to capture history dependent behaviors, such approach may potentially face some limitation when applied to complex microstructures with strongly nonlinear material laws. Some prospects for this work consists on the incorporation of physical based laws in the training process. Recent works by [24,43] has proven that the incorporation of thermodynamic principles in the ANN architecture design or during the training process can significantly improve models reliability.

Nevertheless, all the examples presented in this work constitute a proof of concept of the use of FE-LSTM approaches to accelerate multi-scale simulations of heterogeneous structures. However, it should be kept in mind that many parameters can influence the approach reliability, in particular the microstructure behavior complexity which sometimes may require a lot of data to train the LSTM models. In this study, only octet-truss architected structures with a plastic and isotropic hardening constitutive law were considered without geometric non linearities. The approach has been therefore validated for time-independent plasticity, and requires yet to be validated using other type of material non-linearities (i.e. time dependent responses such as viscoelasticity and viscoplasticity). Finally, it will be interesting in the future to test this approach by considering several architectured materials such as TPMS with their associate constitutive behavior and including the geometric non linearities in the model (which will highly increase the database generation cost and model complexity).

CRediT authorship contribution statement

Aymen Danoun: Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Conceptualization. **Etienne Prulière:** Writing – review & editing, Supervision, Software, Resources, Methodology, Conceptualization. **Yves Chemisky:** Writing – review & editing, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] M.T. Hsieh, M.R. Begley, L. Valdevit, Architected implant designs for long bones: Advantages of minimal surface-based topologies, Mater. Des. 207 (2021) 109838, http://dx.doi.org/10.1016/j.matdes.2021.109838, URL https://www.sciencedirect.com/science/article/pii/S0264127521003919.

[2] F. Feyel, Multiscale FE2 elastoviscoplastic analysis of composite structures, Comput. Mater. Sci. (1999) http://dx.doi.org/10.1016/s0927-0256(99)00077-4.

[3] F. Feyel, A multilevel finite element method (FE^2) to describe the response of highly non-linear structures using generalized continua, Comput. Methods Appl. Mech. Engrg. (2003) http://dx.doi.org/10.1016/S0045-7825(03)00348-7.

[4] E. Tikarrouchine, G. Chatzigeorgiou, F. Praud, B. Piotrowski, Y. Chemisky, F. Meraghni, Three-dimensional FE^2 method for the simulation of non-linear, rate-dependent response of composite structures, Compos. Struct. (2018) http://dx.doi.org/10.1016/j.compstruct.2018.03.072.

[5] E. Tikarrouchine, G. Chatzigeorgiou, Y. Chemisky, F. Meraghni, Fully coupled thermo-viscoplastic analysis of composite structures by means of multi-scale three-dimensional finite element computations, Int. J. Solids Struct. (2019) http://dx.doi.org/10.1016/j.ijsolstr.2019.01.018.

[6] N. Lange, G. Hütter, B. Kiefer, An efficient monolithic solution scheme for FE^2 problems, Comput. Methods Appl. Mech. Engrg. 382 (2021) 113886, http://dx.doi.org/10.1016/J.CMA.2021.113886.

[7] S. Chaouch, J. Yvonnet, An unsupervised machine learning approach to reduce nonlinear FE^2 multiscale calculations using macro clustering, Finite Elem. Anal. Des. 229 (2024) 104069.

[8] N. Lange, G. Hütter, B. Kiefer, A monolithic hyper ROM FE^2 method with clustered training at finite deformations, Comput. Methods Appl. Mech. Engrg. 418 (2024) 116522, http://dx.doi.org/10.1016/j.cma.2023.116522, URL https://www.sciencedirect.com/science/article/pii/S0045782523006461.

[9] H. Moulinec, P. Suquet, A numerical method for computing the overall response of nonlinear composites with complex microstructure, Comput. Methods Appl. Mech. Engrg. 157 (1) (1998) 69–94, http://dx.doi.org/10.1016/S0045-7825(97)00218-1, URL https://www.sciencedirect.com/science/article/pii/S0045782597002181.

[10] J. Michel, P. Suquet, Nonuniform transformation field analysis, Int. J. Solids Struct. 40 (25) (2003) 6937–6955, http://dx.doi.org/10.1016/S0020-7683(03)00346-9, URL https://www.sciencedirect.com/science/article/pii/S0020768303003469 Special issue in Honor of George J. Dvorak.

[11] F. Fritzen, O. Kunc, Two-stage data-driven homogenization for nonlinear solids using a reduced order model, Eur. J. Mech. A Solids 69 (2018) 201–220, http://dx.doi.org/10.1016/j.euromechsol.2017.11.007, URL https://www.sciencedirect.com/science/article/pii/S0997753817306927.

[12] J. Ghaboussi, J.H. Garrett, X. Wu, Knowledge-Based Modeling of Material Behavior with Neural Networks, J. Eng. Mech. 117 (1) (1991) 132–153, http://dx.doi.org/10.1061/(ASCE)0733-9399(1991)117:1(132).

[13] T. Furukawa, G. Yagawa, Implicit constitutive modelling for viscoplasticity using neural networksr, Technical Report, 43, 1998, pp. 195–219, http://dx.doi.org/10.1002/(SICI)1097-0207(19980930)43:2.

[14] S. Jung, J. Ghaboussi, Neural network constitutive model for rate-dependent materials, Comput. Struct. 84 (15–16) (2006) 955–963, http://dx.doi.org/10.1016/J.COMPSTRUC.2006.02.015.

[15] Z. Waszczyszyn, L. Ziemiański, Neural networks in mechanics of structures and materials - New results and prospects of applications, Comput. Struct. 79 (22–25) (2001) 2261–2276, http://dx.doi.org/10.1016/S0045-7949(01)00083-9.

[16] A. Zhang, D. Mohr, Using neural networks to represent von Mises plasticity with isotropic hardening, Int. J. Plast. 132 (2020) http://dx.doi.org/10.1016/j.ijplas.2020.102732.

[17] U. Ali, W. Muhammad, A. Brahme, O. Skiba, K. Inal, Application of artificial neural networks in micromechanics for polycrystalline metals, Int. J. Plast. 120 (May) (2019) 205–219, http://dx.doi.org/10.1016/j.ijplas.2019.05.001.

[18] W.Q. Shen, Y.J. Cao, J.F. Shao, Z.B. Liu, Prediction of plastic yield surface for porous materials by a machine learning approach, Mater. Today Commun. 25 (June) (2020) 101477, http://dx.doi.org/10.1016/j.mtcomm.2020.101477.

[19] C. Settgast, M. Abendroth, M. Kuna, Constitutive modeling of plastic deformation behavior of open-cell foam structures using neural networks, Mech. Mater. 131 (January) (2019) 1–10, http://dx.doi.org/10.1016/j.mechmat.2019.01.015.

[20] R. Ravinder, K.H. Sridhara, S. Bishnoi, H.S. Grover, M. Bauchy, Jayadeva, H. Kodamana, N.M.A. Krishnan, Deep learning aided rational design of oxide glasses, Mater. Horiz. 7 (7) (2020) 1819–1827, http://dx.doi.org/10.1039/D0MH00162G, arXiv:1912.11582, https://pubs.rsc.org/en/content/articlehtml/2020/mh/d0mh00162g https://pubs.rsc.org/en/content/articlelanding/2020/mh/d0mh00162g.

[21] X. Liu, C.E. Athanasiou, N.P. Padture, B.W. Sheldon, H. Gao, A machine learning approach to fracture mechanics problems, Acta Mater. 190 (2020) 105–112, http://dx.doi.org/10.1016/J.ACTAMAT.2020.03.016.

[22] B. Ouyang, Y. Song, Y. Li, F. Wu, H. Yu, Y. Wang, Z. Yin, X. Luo, G. Sant, M. Bauchy, Using machine learning to predict concrete's strength: learning from small datasets, Eng Res Exp 3 (1) (2021) 015022, http://dx.doi.org/10.1088/2631-8695/ABE344, https://iopscience.iop.org/article/10.1088/2631-8695/abe344 https://iopscience.iop.org/article/10.1088/2631-8695/abe344/meta.

[23] L. Benabou, Development of LSTM networks for predicting viscoplasticity with effects of deformation, strain rate, and temperature history, J. Appl. Mech. 88 (7) (2021) 071008.

[24] A. Danoun, E. Prulière, Y. Chemisky, Thermodynamically consistent Recurrent Neural Networks to predict non linear behaviors of dissipative materials subjected to non-proportional loading paths, Mech. Mater. 173 (2022) 104436, http://dx.doi.org/10.1016/j.mechmat.2022.104436, URL https://www.sciencedirect.com/science/article/pii/S0167663622002034.

[25] S. Khandelwal, S. Basu, A. Patra, A machine learning-based surrogate modeling framework for predicting the history-dependent deformation of dual phase microstructures, Mater. Today Commun. 29 (2021) 102914, http://dx.doi.org/10.1016/j.mtcomm.2021.102914, URL https://www.sciencedirect.com/science/article/pii/S2352492821009016.

[26] S. Im, J. Lee, M. Cho, Surrogate modeling of elasto-plastic problems via long short-term memory neural networks and proper orthogonal decomposition, Comput. Methods Appl. Mech. Engrg. 385 (2021) 114030, http://dx.doi.org/10.1016/j.cma.2021.114030, URL https://www.sciencedirect.com/science/article/pii/S0045782521003613.

[27] W.G. Dettmer, E.J. Muttio, R. Alhayki, D. Perić, A framework for neural network based constitutive modelling of inelastic materials, Comput. Methods Appl. Mech. Engrg. 420 (2024) 116672, http://dx.doi.org/10.1016/j.cma.2023.116672, URL https://www.sciencedirect.com/science/article/pii/S0045782523007958.

[28] B.A. Le, J. Yvonnet, Q.C. He, Computational homogenization of nonlinear elastic materials using neural networks, Internat. J. Numer. Methods Engrg. 104 (12) (2015) 1061–1084, http://dx.doi.org/10.1002/NME.4953, https://onlinelibrary.wiley.com/doi/full/10.1002/nme.4953 https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.4953 https://onlinelibrary.wiley.com/doi/10.1002/nme.4953.

[29] V. Minh Nguyen-Thanh, L. Trong Khiem Nguyen, T. Rabczuk, X. Zhuang, A surrogate model for computational homogenization of elastostatics at finite strain using high-dimensional model representation-based neural network, Internat. J. Numer. Methods Engrg. 121 (21) (2020) 4811–4842, http://dx.doi.org/10.1002/NME.6493, https://onlinelibrary.wiley.com/doi/full/10.1002/nme.6493 https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.6493 https://onlinelibrary.wiley.com/doi/10.1002/nme.6493.

[30] X. Lu, D.G. Giovanis, J. Yvonnet, V. Papadopoulos, F. Detrez, J. Bai, A data-driven computational homogenization method based on neural networks for the nonlinear anisotropic electrical response of graphene/polymer nanocomposites, Comput. Mech. 64 (2) (2018) 307–321, http://dx.doi.org/10.1007/S00466-018-1643-0, URL https://link.springer.com/article/10.1007/s00466-018-1643-0.

[31] L.C. Jain, L.R. Medsker, Recurrent Neural Networks: Design and Applications, first ed., CRC Press Inc., USA, 1999.

[32] M.B. Gorji, M. Mozaffar, J.N. Heidenreich, J. Cao, D. Mohr, On the potential of recurrent neural networks for modeling path dependent plasticity, J. Mech. Phys. Solids (2020) http://dx.doi.org/10.1016/j.jmps.2020.103972.

[33] D.W. Abueidda, S. Koric, N.A. Sobh, H. Sehitoglu, Deep learning for plasticity and thermo-viscoplasticity, Int. J. Plast. 136 (2021) 102852, http://dx.doi.org/10.1016/j.ijplas.2020.102852.

[34] P. Suquet, Elements of homogenization for inelastic solid mechanics, Lecture Notes in Phys. 272 (1987) 193–278.

[35] G. Chatzigeorgiou, N. Charalambakis, Y. Chemisky, F. Meraghni, Periodic homogenization for fully coupled thermomechanical modeling of dissipative generalized standard materials, Int. J. Plast. 81 (2016) 18–39, http://dx.doi.org/10.1016/j.ijplas.2016.01.013, URL https://www.sciencedirect.com/science/article/pii/S0749641916300043.

[36] S. Hochreiter, J. Schmidhuber, Long Short-Term Memory, Neural Comput. 9 (8) (1997) 1735–1780, http://dx.doi.org/10.1162/neco.1997.9.8.1735, URL http://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf.

[37] P.J. Werbos, Backpropagation through time: What it does and how to do it, Proc. IEEE 78 (10) (1990) 1550–1560, http://dx.doi.org/10.1109/5.58337.

[38] E. Prulière, Y. Chemisky, 3MAH : un ensemble de librairies pour analyser le comportement complexe de matériaux hétérogènes, 2022.

[39] F. Chollet, et al., Keras, 2015, GitHub, https://github.com/fchollet/keras.

[40] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-scale machine learning on heterogeneous systems, 2015, URL https://www.tensorflow.org/ Software available from tensorflow.org.

[41] S.J. Prince", Understanding Deep Learning, The MIT Press, 2023.

[42] F. Fritzen, M. Fernández, F. Larsson, On-the-fly adaptivity for nonlinear twoscale simulations using artificial neural networks and reduced order modeling, Front. Mater. 6 (2019) http://dx.doi.org/10.3389/fmats.2019.00075.

[43] F. Masi, I. Stefanou, P. Vannucci, V. Maffi-Berthier, Thermodynamics-based artificial neural networks for constitutive modeling, J. Mech. Phys. Solids 147 (2021) 104277, http://dx.doi.org/10.1016/J.JMPS.2020.104277, arXiv:2005.12183.